NON-UNIFORM SAMPLING: ALGORITHMS AND ARCHITECTURES

A Thesis
Presented to
The Academic Faculty

by

Chenchi Luo

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
December 2012

NON-UNIFORM SAMPLING: ALGORITHMS AND ARCHITECTURES

Approved by:

Professor Justin K. Romberg,
Committee Chair
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor James H. McClellan,
Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor David V. Anderson
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Mark A. Davenport
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Santosh S. Vempala
College of Computing
Georgia Institute of Technology

Date Approved: 8 November 2012

This dissertation is dedicated to my parents,

Zhenyuan Luo and Binhua Xu,

Thanks for your love and support.

ACKNOWLEDGEMENTS

In retrospect to the long journey I spent with Georgia Tech, I can never forget the support, encouragement, and help from people who accompanied me during this adventure. I would like to take this opportunity to express my heartfelt gratitude to all of them at this time.

The first two persons that I would like to thank is my father, Zhenyuan Luo and my mother, Binhua Xu. They have been a consistent driving force for me to exploit my potentials and to strive for excellence. They made my life their priority and put their own careers behind and unconditionally supported me in the decision of pursuing a Ph.D degree both financially and emotionally. When I got frustrated, it was their love and encouragement that kept me sticking to my goal and never giving up.

The next person I'd like to thank is my advisor Dr. James H. McClellan. I simply cannot expect a better advisor. I was amazed by his profound knowledge both inside and outside the specialty in signal processing. He is my mentor not only in passing the knowledge and expertise but also in setting a model as what kind of person I want to be. Dr. McClellan is a real master in eduction. He can always explain the most complicated concepts in a simple and vivid language. Without his illuminating direction I could end up with a lot of detours in my research. His supervision style gave me a high degree of freedom in forming my own path in pursuit of the research topics that I was interested in.

Another faculty member that I'd like to present my gratitude to is Dr. Guotong Zhou, who initiated the GT-Shanghai Dual Master Program collaborated with Shanghai Jiao Tong University. I was lucky and proud to be among the first few students

that joined this program in 2006. This program entitled me a chance to experience the courses taught by known professors and functioned as a corner stone for my pursuit of a Ph.D. degree later. I'd also like to thank Dr. Xiaoli Ma, Dr. G.K. Chang, Dr. Monty. M. Hayes from Georgia Tech who traveled all the way to Shanghai to gave us those excellent lectures. I'm also grateful to the members of my thesis proposal and defense committee, Dr. Justin Romberg, Dr. David V. Anderson, Dr. William D. Hunt, who provided valuable feedbacks in improving my existing work and exploring new fields.

I'd also like to thank Texas Instruments Leadership University program that sponsored my research. I'm especially grateful to Dr. Arthur J. Redfern and Dr. Milind. A. Borkar who supervised me in the summer of 2010 and 2011 when I was an intern at TI System and Applications R&D Center. Those two fruitful summers led to the production of multiple publications and patents.

Finally, I would like to thank my friends, Renke Huang, Jie Tan, Chao Wang, who I know since the first time I entered college back in 2002. It's such an honor to spend the most valuable decade of my life with them going through the same university both at home and abroad, escalating from a B.S. to Ph.D. I'd also like to thank Yuting Gu, Zhenxian Wang, Hong Yu and the members in my research group: Kyle Krueger, Carson Wick, Lingchen Zhu, Wei-Hsin Mu, Peter Tuuk and Gregory Krudysz.

Contents

List of Tables

List of Figures

SUMMARY

Modern signal processing applications emerging in telecommunication and instrumentation industries have placed an increasing demand for ADCs with higher speed and resolution. The most fundamental challenge in such a progress lies at the heart of the classic signal processing: the Shannon-Nyquist sampling theorem which stated that when sampled uniformly, there was no way to increase the upper frequency in the signal spectrum and still unambiguously represent the signal except by raising the sampling rate. This thesis is dedicated to the exploration of the ways to break through the Shannon-Nyquist sampling rate by applying non-uniform sampling techniques.

Time interleaving is probably the most intuitive way to parallel the uniform sampling process in order to achieve a higher sampling rate. Unfortunately, the channel mismatches in the TIADC system make the system an instance of a recurrent non-uniform sampling system whose non-uniformities are detrimental to the performance of the system and need to be calibrated. Accordingly, this thesis proposes a flexible and efficient architecture to compensate for the channel mismatches in the TIADC system. As a key building block in the calibration architecture, the design of the Farrow structured adjustable fractional delay filter has been investigated in detail. A new modified Farrow structure is proposed to design the adjustable FD filters that are optimized for a given range of bandwidth and fractional delays. The application of the Farrow structure is not limited to the design of adjustable fractional delay filters. It can also be used to implement adjustable lowpass, highpass and bandpass filters as well as adjustable multirate filters. This thesis further extends the Farrow structure to the design of filters with adjustable polynomial phase responses.

Inspired by the theory of compressive sensing, another contribution of this thesis is to use randomization as a means to overcome the limit of the Nyquist rate. This thesis investigates the impact of random sampling intervals or jitters on the power spectrum of the sampled signal. It shows that the aliases of the original signal can be well shaped by choosing an appropriate probability distribution of the sampling intervals or jitters such that aliases can be viewed as a source of noise in the signal power spectrum. A new theoretical framework has been established to associate the probability mass function of the random sampling intervals or jitters with the aliasing shaping effect. Based on the theoretical framework, this thesis proposes three random sampling architectures, i.e., SAR ADC, ramp ADC and level crossing ADC, that can be easily implemented based on the corresponding standard ADC architectures. Detailed models and simulations are established to verify the effectiveness of the proposed architectures. A new reconstruction algorithm called the successive sine matching pursuit has also been proposed to recover a class of spectrally sparse signals from a sparse set of non-uniform samples onto a denser uniform time grid so that classic signal processing techniques can be applied afterwards.

Chapter I

INTRODUCTION

The real world is analog in nature. On the other hand, computers are digital. Analog to digital converters (ADCs) are responsible for filling the gap between the real world and computers. Only after being converted to digital forms can the original analog signals be further processed by computers for such purposes as estimation, detection and filtering. Modern signal processing applications emerging in telecommunication and instrumentation industries have placed an increasing demand for ADCs with higher speed and resolution. The most fundamental challenge in such a progress lies at the heart of the classic signal processing: the Shannon sampling theorem [69], which states that a band-limited analog signal that has been periodically sampled can be perfectly reconstructed from an infinite sequence of samples if the sampling rate exceeds $2B$ samples per second (the Nyquist rate), where B is the highest frequency in the original signal. In other words, when uniform or periodic sampling scheme is adopted, there is no way to increase the upper frequency in the signal spectrum and still unambiguously represent the signal except by raising the sampling rate. Therefore, conventional ADC manufacturing technologies dictate the highest frequencies that can be handled digitally. Although the microelectronic device manufacturing technologies are continuously improving, the upper frequency limit in the digital domain is increasing at a relatively slow pace.

As semiconductor manufacturing technologies reach their limits, central processing units (CPUs) in computers have resorted to parallelism to circumvent the obstacle. CPUs today are manufactured with multi-cores so that tasks can be processed in a parallel manner instead of sequentially. Similarly, ADCs can also be arranged in

a parallel way so that a group of slower ADCs take samples alternately and the outputs of all channels are multiplexed to achieve a faster equivalent sampling rate. This kind of structure is referred to as time-interleaved ADC (TIADC) architecture in literatures. The concept of time-interleaving is not restricted to increasing the equivalent sampling speed only. It has also been applied to oversampling ADCs [40, 41] to reach a higher resolution instead of speed. Ideally, the sampling speed would scale linearly with the number of channels. However, channel mismatches ultimately limit the performance of time-interleaved ADCs. From a theoretical perspective, the time-interleaved ADC system can be viewed as a special case of recurrent non-uniform sampling. The objective of calibration is to interpolate the non-uniformly sampled signal onto a uniform time grid. The conventional way to identify and compensate channel mismatches is to calibrate the ADCs before use. The problem with offline calibration is that the mismatches usually vary during the lifetime of ADCs due to temperature change or device aging. Therefore, one contribution of the thesis is the proposal of a flexible and efficient architecture to re-calibrate the time-interleaved ADC system without the interruption of the data conversion process.

Among various kinds of channel mismatches, the calibration of timing mismatches is probably the most demanding one to deal with. This thesis conducts a comprehensive study on the design of digital fractional delay (FD) filters, especially a filter structure called the Farrow structure in the implementation of FD filters in order to calibrate the timing mismatches online. The second contribution of the thesis is a new modified Farrow structured FD filter whose performance is shown to be superior to the existing design methods.

As a byproduct, the third contribution of the thesis is the extension of the Farrow structure in the implementation of scalable multi-rate filter, adjustable bandwidth filter and adjustable polynomial phase filter structures.

The above mentioned time-interleaved ADC system still satisfies the Nyquist criterion as the equivalent sampling frequency of the TIADC system is still at or above the Nyquist frequency of the signal. The non-uniformities imbedded in the system are introduced passively, thus are unknown and need to be compensated. The multi-coset sampling system [64, 87, 9] also belongs to the recurrent non-uniform sampling system. It is similar to TIADC system except for two differences. First, the equivalent sampling frequency of the multi-coset sampling system can be set below the Nyquist frequency of the signal. Second, the timing mismatches are precisely controllable. It is shown that we can still perfectly recover a class of sparse multi-band signal even sampling below the Nyquist frequency.

Another strategy to overcome the limit of the Nyquist rate is to jump out of the uniform sampling constraint. The Nyquist criterion is a sufficient condition for lossless and unique signal acquisition but not a necessary one. Non-uniformities can be introduced purposefully in the sampling process for signal preservation and reconstruction. Compressive sensing theory [12] states that for a class of sparse signals, sampling can be achieved at a rate slightly higher than the information rate in the signal rather than its Nyquist rate. Randomization is deliberately introduced as a tool in the signal conversion operations, including both sampling time and amplitude quantization. Randomization means that there is no longer periodicity in the sampling process, which suggests the name of non-recurrent non-uniform sampling systems. The fourth contribution of this thesis is the investigation of the impact of the probabilistic distributions of the sampling time points on the spectra of the non-uniformly sampled signals. A novel perspective is proposed to treat the aliases as a noise source introduced by the non-uniform sampling process. Accordingly, we can shape the power distribution of the aliasing noise over the signal spectrum by designing an appropriate non-uniform sampling process so that the signal spectrum remains detectable in the presence of the aliasing noise floor. A new reconstruction

algorithm is proposed to reconstruct a class of spectrally sparse signals from a sparse and non-uniform time grid onto a dense and uniform time grid.

Inspired by the theory of compressive sensing, numerous innovative analog to digital converter architectures have been proposed, among which the random demodulator [76] and modulated wideband converter [53, 51, 52] are the two most prominent architectures. However, these systems require specialized hardware such as analog mixing circuits, pseudo random number generator and filters, the cost of which is not affordable in many applications. The fifth contribution of the thesis is to apply non-uniform sampling to standard ADC architectures such as successive approximation register (SAR) ADCs, ramp ADCs and level crossing (LC) ADCs and show that a reconstruction of the signal is possible even when it is sampled below the Nquist frequency.

1.1 Non-uniform Sampling and Reconstruction

In 1949, Shannon published the paper "Communication in the Presence of Noise [69]", which set the foundation of information theory. According to the Shannon-Nquist sampling theorem, a band-limited signal $x(t)$ can be exactly reconstructed from its uniform samples $x[nT_s]$. The perfect reconstruction formula is given by

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\,\mathrm{sinc}(2B(t - nT_s)), \tag{1}$$

where $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$ and $x(t)$ is band-limited to $[-B, B]$. The inter-sample time interval is given by $T_s = 1/2B$. Reconstruction is achieved by convolving the uniform samples with a sinc function, which is equivalent to an ideal low-pass filtering (LPF) process in frequency domain.

Shannon's sampling theorem is more than 60 years old. Thanks in part to the mathematical connections that were made with the wavelet theory, there has been a revival of this topic and a lot of new findings were made in the past two decades.

Jerri [31] and Unser [78] gave two comprehensive overviews of various extensions and applications of the classic sampling theorem.

Higgins [30] generalized Shannon's sampling theorem to non-uniformly spaced samples. Denote t_n as the sampling time instance, Higgins' theorem states that suppose $\{e^{j2\pi f t_n}\}$ is a complete basis for $X(f)$ within the bandwidth $[-B, B]$, then the functions $\{g(t, t_n)\} = \text{sinc}(2B(t - t_n))$ is also a complete basis for $x(t), -\infty < t < \infty$. There will always exist a unique basis $\{k(t, t_n)\}$ that is bi-orthogonal to $\{g(t, t_n)\}$. The bi-orthogonality means that the elements within the respective bases are not orthogonal, but are orthogonal to the elements of the other basis.

$$< g(t, t_n), k(t, t_n) >= \int_{n=-\infty}^{\infty} g(t, t_n), k(t, t_n) dt = \delta_{mn}. \tag{2}$$

We can then represent $x(t)$ in two ways:

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{\infty} < x(u), k(u, t_n) > g(t, t_n) \\ &= \sum_{n=-\infty}^{\infty} c_n g(t, t_n), \end{aligned} \tag{3}$$

$$\begin{aligned} x(t) &= \sum_{n=-\infty}^{\infty} < x(u), g(u, t_n) > k(t, t_n) \\ &= \sum_{n=-\infty}^{\infty} x(t_n) k(t, t_n). \end{aligned} \tag{4}$$

The prerequisite of the theorem raises a question as what set of $\{t_n\}$ makes $\{e^{j2\pi f t_n}\}$ a complete basis for $X(f)$. Kadec [36] proved that this was the case when

$$|t_n - nT_s| < 1/4T_s. \tag{5}$$

Beutler [3] found that $\{e^{j2\pi f t_n}\}$ was a complete basis when the average sampling rate exceeded the Nyquist rate of the signal. Another problem is how to evaluate $\{k(t, t_n)\}$. Generally speaking, given a set of $\{t_n\}$, there is no closed form expression

for $\{k(t,t_n)\}$. However, Higgins showed that if (5) is satisfied, $\{k(t,t_n)\}$ is a Lagrange interpolation function given by

$$k(t,t_n) = \frac{H(t)}{H'(t_n)(t-t_n)}, \tag{6}$$
$$H(t) = (t-t_0)\prod_{m\neq 0}(1-t/t_m).$$

Under uniform sampling, the Lagrange interpolation function degenerates into an ideal linear time-invariant low-pass filter. While under non-uniform sampling, the Lagrange interpolation function is time-variant and has no closed form expression, which makes it infeasible in many practical real-time applications. Various researches have studied and compared simpler forms of interpolation functions such as sample-and-hold, n-th order hold [50] and spline functions [79, 80]. Yen [94] provided a minimum mean squares error estimate of the Lagrange interpolation function $\phi_n(t)$ based on a finite set of N samples.

$$\phi_n(t) = \sum_{m=1}^{N} \alpha_{m,n}\,\mathrm{sinc}(2B(t-t_m)) \quad m,n = 0,...,N-1, \tag{7}$$

where $\alpha_{m,n}$ is an element in the inverse of a matrix whose element is given by $\mathrm{sinc}(2B(t_n - t_m))$.

A more general form of the reconstruction problem on a finite sample set of size N is given by

$$x(t_n) = \sum_{m=0}^{M-1} c_m \phi_m(t_n), \quad 0 \le n \le N-1, \tag{8}$$

where the basis function $\phi_n(t)$ is either orthogonal or a frame, which is complete but not necessarily orthogonal. Then it is necessary to find the coefficients c_m, which involves solving an M by N system of linear equations. The singular value decomposition (SVD) [93] provides a direct method to solve the problem. An iterative method [49, 14] was proved to be a more efficient way to find the pseudo-inverse of the basis matrix.

1.2 Time-Interleaved ADC System

Another remarkable finding in the realm of sampling theory was made by Papoulis [58] who proposed a generalized sampling expansion in 1977. He showed that a band-limited signal $x(t)$ can be uniquely determined by the samples $g_k(nT_s)$, which are the outputs of M linear systems $g_k(t)$ with input $x(t)$, sampled at $1/M$ times the Nyquist rate. One example of such a system is to sample $x(t)$ and its m-th order derivatives $x^{(m)}(t), (m = 1, ...M - 1)$ at $1/M$ the Nyquist rate. While the generalized sampling concept is straightforward, Papoulis did not investigate the reconstruction process in terms of stability and causality [16, 11]. The perfect reconstruction for maximally decimated filter banks [81] can be viewed as a discrete time counter part of the generalized sampling expansion. In 1981, Brown [10] discovered the link between generalized sampling and multi-rate system and proposed a multi-channel sampling structure for band-limited signals. The mixing of continuous-time and discrete-time filters in the sub-channels gave rise to the study of hybrid filter banks (HFBs) [82, 63, 86, 71, 22, 56, 45]. The HFBs are multi-rate systems that can be decomposed into three stages: an analog analysis filter bank, A/D conversion, and a digital synthesis filter bank.

Time-interleaved ADCs can be viewed as a special form of the hybrid filter banks. Since the ADCs in each channel are not exactly identical due to discrepancies in the manufacturing process, channel mismatches in time-interleaved ADC system make the system a recurrent non-uniform sampling process. The DC offset, gain and timing mismatches in the interleaved ADC system were initially investigated under the framework of HFBs by Petraglia [62]. A unified and more general treatment can be found in [88, 90]. The non-linear behaviors of ADCs were modeled by exploiting non-linear HFBs in [91, 89].

1.3 Random Sampling Theory

Digital aliasing-free signal processing (DASP) was first mentioned by Shapiro and Silverman [70] in 1960. The key idea is to randomize the placement of sampling points in order to suppress or eliminate aliasing. Beutler and Leneman [5, 4, 46, 6] published a series of papers on the theory of stationary point process and random sampling of random process in the late 1960s. Bilinskis [8] later showed that aliasing free sampling is theoretically possible whenever the random sampling point process is stationary. In other words, deterministic sampling is never aliasing free. Given a finite number of samples, aliasing can only be suppressed but not completely eliminated. A comprehensive coverage of topics in DASP, especially a variety of randomization techniques in the sampling process can be found in Bilinskis' book [7] published in 2007.

1.4 Compressive Sensing Theory

Since the original paper by Candes, Romberg, and Tao [12], the signal processing community has been overwhelmed by the possibilities offered by compressive sensing (CS). CS has inspired hundreds of papers in the fields of medical imaging [48], array signal processing [27], geoscience [28] and telecommunication [2]. Two excellent tutorials about CS can be found in [13] and [1].

CS concerns itself with the scenario where a signal $\mathbf{x} \in \mathbb{C}^N$ is sparse in a known basis $\mathbf{\Psi} \in \mathbb{C}^{N \times N}$. The representation of $\mathbf{x}$ under $\mathbf{\Psi}$ involves a sparse vector $\boldsymbol{\alpha}$ with a sparsity K

$$\mathbf{x} = \mathbf{\Psi}\boldsymbol{\alpha}, \tag{9}$$

where there are at most $K (K \ll N)$ non-zero entries in $\boldsymbol{\alpha}$.

Instead of observing the signal $\mathbf{x}$ itself, we observe the signal $\mathbf{y} \in \mathbb{C}^M$ through a projection matrix $\mathbf{\Phi} \in \mathbb{C}^{M \times N}$.

$$\mathbf{y} = \mathbf{\Phi}\mathbf{x} = \mathbf{\Phi}\mathbf{\Psi}\boldsymbol{\alpha} = \mathbf{A}\boldsymbol{\alpha}, \tag{10}$$

where $\mathbf{A}$ is the measurement operator.

$\mathbf{A}$ satisfies a so called restricted isometry property (RIP) with constant δ_K if

$$\sqrt{1-\delta_K}||\boldsymbol{\alpha}||_2 \leq ||\mathbf{A}\boldsymbol{\alpha}||_2 \leq \sqrt{1+\delta_K}||\boldsymbol{\alpha}||_2, \quad \forall \boldsymbol{\alpha} \text{ with } ||\boldsymbol{\alpha}||_0 \leq K. \tag{11}$$

RIP states that the vector norm of any K sparse vector is preserved when operated on by $\mathbf{A}$, which makes it possible to recover $\mathbf{x}$ from its low dimension embedding $\mathbf{y}$. It is a non-deterministic polynomial-time (NP) hard problem to verify that a matrix that satisfies the RIP property. To overcome this obstacle, researchers have embraced random matrices which can be shown to satisfy the RIP property with very high probability. In particular, if each entry of $\mathbf{A}$ is drawn according to a sub-gaussian probability distribution, the RIP will be satisfied with high probability as long as $M = \mathcal{O}(K\log(N/K))$. In terms of sampling, this implies that the minimum sampling rate is roughly proportional to the information rate in the signal.

Once the compressed measurement vector $\mathbf{y}$ is acquired, the next objective is to reconstruct $\mathbf{x}$ according to

$$\min_{\mathbf{x}} ||\mathbf{x}||_0, \quad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{12}$$

Unfortunately, (12) is non-convex as well as NP hard. The RIP property of $\mathbf{A}$ enables a convex relaxation of (12) that yields the same solution as (12)

$$\min_{\mathbf{x}} ||\mathbf{x}||_1, \quad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}. \tag{13}$$

When measurement noise is present, (13) is reformulated as

$$\min_{\mathbf{x}} ||\mathbf{x}||_1, \quad \text{s.t. } ||\mathbf{y} = \mathbf{A}\mathbf{x}||_2 \leq \epsilon. \tag{14}$$

In the CS community, (13) and (14) are referred as basis pursuit (BP) and basis pursuit de-noising (BPDN) problems, respectively. A large variety of algorithms have been proposed to solve these problems. Optimization based solvers [42] were among the first ones to be investigated. Despite their accuracy and convergence guarantee,

these methods are slow in general, thus not suitable for large scale problems. Alternatively, greedy methods [54] were used to give an approximate solution to (13) and (14). Orthogonal matching pursuit (OMP) [75] was initially proposed by Tropp et al, and several variants of OMP were invented to further improve the performance of OMP. Among those variants, compressive sampling matching pursuit (CoSaMP) [55] offered the most attractive guarantee for convergence and computational complexity. Although a rich choice of algorithms exists, the selection of an appropriate recovery algorithm, which depends on available computation resources and the nature of the problem is problem-specific.

1.4.1 Compressive Signal Acquisition

CS is a fundamentally discrete problem. Attempts were made to apply CS to acquire continuous time signals compressively. Two promising methods based on random analog filtering of the input signal followed by uniform sub-Nyquist rate sampling were proposed. A random demodulator converter architecture was proposed by Tropp et al. [76, 74]. Figure 1 shows the basic architecture of the proposed converter. First, the input signal $x(t)$ is modulated by a pre-determined Bernoulli distributed random sequence $p(t)$ whose chipping rate equals the Nyquist rate of the signal. Therefore, $P(f)$ has a spread spectrum over the Nyquist range. In the frequency domain, $X(f)$ is convolved with $P(f)$ so that the high frequency components in $X(f)$ are modulated into the baseband of $\tilde{X}(f)$. After low-pass filtering, the wideband information in $X(f)$ is still preserved. The final step is a sub-Nyquist rate uniform sampler. A corresponding recovery algorithm was proposed to reconstruct $x(t)$ onto a uniform time grid at the Nyquist rate. A prerequisite for this architecture to work is that $x(t)$ is a sparse multi-sine signal which has only a small number of nonzero coefficients in its Fourier series expansion. Besides, the random demodulator architecture also requires that the frequencies of $x(t)$ lies on the DFT grid. For those frequencies

lying off the grid, perfect recovery cannot be achieved due to spectral leakage. As

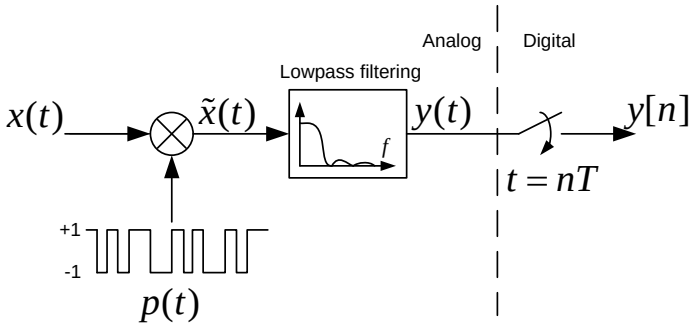

Figure 1: Block diagram of the random demodulator converter architecture.

an alternative, a so-called modulated wideband converter (MWC) architecture was proposed by Mishali and Eldar [51]. The MWC is composed of M channels, each channel is similar to a random demodulator except that in the random demodulator architecture, the timing function used for lowpass filtering is a rectangular window while it is a sinc function in the MWC. The input signal assumption of the MWC is also different from the random demodulator. The MWC assumes a multi-band signal model consisting of a small number of band-limited signals, each with a continuous frequency support. The number of channels determines the number of signal bands that can be acquired. Although the MWC is more general and robust than the random demodulator, extra channels are needed and the signal reconstruction part is more involved.

1.4.2 Fourier Random Sampling

The Fourier random sampling problem [26] which was well studied before the emergence of CS can be stated in the following way. Given a signal $\mathbf{x} \in \mathbb{C}^N$, find an optimal Fourier representation $\mathbf{x}_{\text{opt}}$ of K complex exponential terms to approximate $\mathbf{x}$. This clearly can be done by performing the fast Fourier transform (FFT) of $\mathbf{x}$ and locating the K largest terms. Gilbert et al. showed in [26] that we can find a Fourier

representation $\mathbf{x}^*$ by only sampling a subset $T \subseteq [0, N-1]$ of $\mathbf{x}$ such that

$$||\mathbf{x} - \mathbf{x}^*||_2^2 \leq (1+\epsilon)||\mathbf{x} - \mathbf{x}_{\text{opt}}||_2^2, \tag{15}$$

where ϵ is an error bound parameter.

The cardinality of T is polynomial in $K\log(N)\log(M)/\epsilon$, where M is the ratio of the largest to smallest numerical quantity encountered. Therefore, the number of samples is sub-linear in N. The subset T can be created by conducting independent Bernoulli trials on the index set $[0, N-1]$, meaning each element of $\mathbf{x}$ has an equal probability of being selected.

In the language of compressive sensing, we can express $\mathbf{x}$ as

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{K-1} \alpha_k e^{j2\pi\omega_k n/N}, \quad \omega_k \subseteq [0, N-1]. \tag{16}$$

Written in matrix form,

$$\mathbf{x} = \mathbf{F}\boldsymbol{\alpha}, \tag{17}$$

where the elements of the discrete time Fourier (DFT) matrix $\mathbf{F}$ are given by $F_{\omega,t} = \frac{1}{\sqrt{N}} e^{j2\pi\omega t/N}$, $\omega, t = 0, ..., N-1$, and $\boldsymbol{\alpha}$ only has K non-zero values at frequencies ω_k. The objective is to obtain $\boldsymbol{\alpha}$ from random samples of $\mathbf{x}$. The random sample set is constructed by running an independent Bernoulli test for each sample. Each sample is selected with probability M/N or discarded with probability $1 - M/N$ so that the expected size of the random sample set is M. Rudelson and Vershynin [65] showed that we can recover the sparse vector $\boldsymbol{\alpha}$ with high probability if $M = \mathcal{O}(K\log^4 N)$.

The Fourier random sampling problem suggests a signal acquisition architecture that can be built upon standard uniform sampling ADCs. At each uniform sampling time point, a Bernoulli random number generator decides whether to sample the signal at the current time point or not according to a prescribed probability. After acquiring M samples, a reconstruction algorithm will be executed to recover the signal on the un-sampled time points. Since less samples are taken per unit time,

this architecture consumes less power. The system complexity has shifted from the analog data acquisition end to the digital reconstruction end.

1.5 Organization of the Thesis

The thesis is organized as follows:

Chapter 2 is dedicated to the discussion of recurrent non-uniform sampling systems. Section 2.1 analyzes the impact of channel mismatches on the output spectrum of the TIADC system. Section 2.2 proposes an adaptive TIADC mismatch calibration algorithm and the corresponding architecture. Section 2.3 deals with the critical building block in the calibration architecture – the design and implementation of digital fractional delay filters, together with a proposal for further improvement and innovations. Section 2.4 discusses another sampling scheme in the recurrent non-uniform sampling category – the multi-coset sampling system, where the spectral sparsity property is utilized to make sub-Nyquist frequency sampling and recovery possible.

Chapter 3 is dedicated to the discussion of non-recurrent non-uniform sampling systems. Section 3.1 reviews the random sampling theory in the field of digital aliasing-free signal processing and connect it with the compressive sensing theory. Section 3.2 analyzes the effect of time quantization, where the randomly distributed sampling intervals or jitters are quantized with a fixed granularity. It is shown that we can shape the power distribution of the aliases by selecting an appropriate sampling scheme to minimize the influence of aliases on the signal spectrum. Section 3.3 proposes several non-recurrent non-uniform sampling architectures that are built upon the existing standard ADC architectures. A new reconstruction algorithm is also proposed to recover a class of spectrally sparse signals from a sparse and non-uniform time grid to a dense and uniform time grid.

Chapter 4 summarizes the key contributions of this thesis and discusses future

directions for this research.

Chapter II

RECURRENT NON-UNIFORM SAMPLING

This chapter investigates two instances of recurrent non-uniform sampling: the time interleaved ADC system and the multi-coset sampling system. The recurrency lies in the fact that the sampling patterns of both systems repeat themselves periodically. In the TIADC system, each channel samples at a rate below the Nyquist rate of the input signal. However, the equivalent sampling rate of the overall system is still above the Nyquist rate of the signal. The non-uniformities are introduced passively due to the discrepancies of each channel. The impact of channel mismatches in the TIADC system is analyzed in Section 2.1. An adaptive TIADC mismatch calibration algorithm and the corresponding architecture is proposed in Section 2.2 and the design of a critical component in the calibration architecture – the digital fractional delay filter is covered in Section 2.3. In the multi-coset sampling system introduced in Section 2.4, the timing mismatches are assumed to be precisely controllable and the input signal is assumed to have a low spectrum occupancy over a wide frequency band. As a result, the equivalent overall sampling rate can be set below the Nyquist rate of the signal while still guarantees a perfect recovery of the input signal without aliasing.

2.1 TIADC System Modeling by Means of Non-linear Hybrid Filter Bank

To understand the non-uniformities in the sampling process of a time-interleaved ADC system, it is essential to establish a detailed model of the TIADC system that reveals the impact of channel mismatches on distorting the output spectrum.

Figure 2 shows the basic structure of an ideal TIADC system. The input analog

signal is sampled in each channel at a rate of F_s/M, where F_s is the Nyquist frequency of $x(t)$, M is the number of channels. The phase of each channel sampler is separated by $2\pi/M$. The digital output of each channel is multiplexed to yield a stream of digital output $y[n]$ at an equivalent sampling frequency of F_s. Figure 3 shows an example of the outputs of a 4 channel TIADC system.

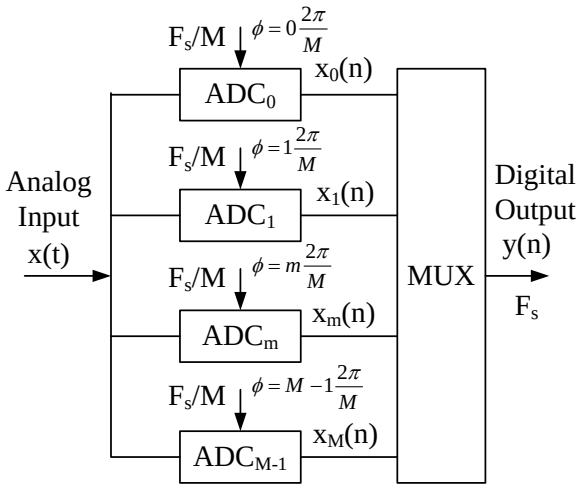

Figure 2: An ideal TIADC system.

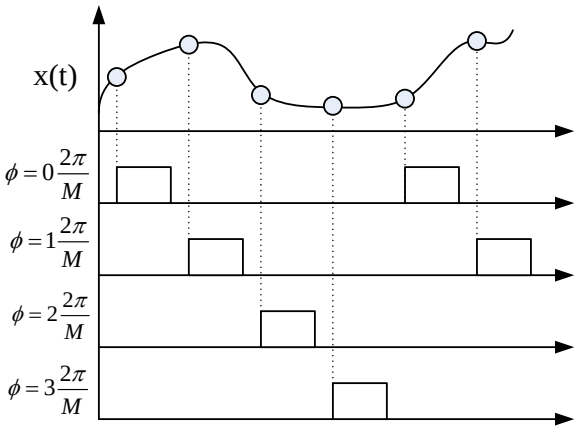

Figure 3: The outputs of an ideal 4 channel TIADC system.

However, in practice, the ADCs in each channel are not identical due to manufacturing discrepancies. Figure 4 shows a more practical ADC model for channel m. The analog input signal $x(t)$ is captured by a track and hold (T/H) circuit whose transfer function can be expressed as $H_m^{TH}(j\Omega)$. The output $x_m^{TH}(t)$ is then sampled

at time $t = (nM + m)T_s$, where $m = 0, ..., M-1$ is the channel index, $n = 0, ..., \infty$ is the time index, $T_s = 1/F_s$. There is a static time delay θ_m^S between the clock signal triggering the sampling event and the actual time samples are taken. Besides the static time delay, there is also a random timing jitter $\theta_m^R[n]$ that can only be described statistically and is both channel and time dependent. The sampled signal can be expressed as

$$x_m^S[n] = x_m^{TH}((nM + m)T_s + \theta_m^S + \theta_m^R[n]). \tag{18}$$

Next, $x_m^S[n]$ is mapped by the transfer function (TF) $f_m(x)$ of the ADC in channel m to yield $x_m^{TF}[n]$ and finally quantized to a digital output $x_m[n]$.

Suppose the transfer function $f_m(x)$ is a memoryless polynomial of order $K-1$:

$$f_m(x) = c_{0,m} + c_{1,m}x + ...c_{K-1,m}x^{K-1}, \tag{19}$$

we can exchange the sequence of the transfer function and sampler, resulting in a model shown in Figure 5. The T/H circuit and the phase advance element can be combined into an analog analysis filter

$$H_m(j\Omega) = H_m^{TH}(j\Omega)e^{j\Omega(\theta_m^S + \theta_m^R[n] + mT_s)}. \tag{20}$$

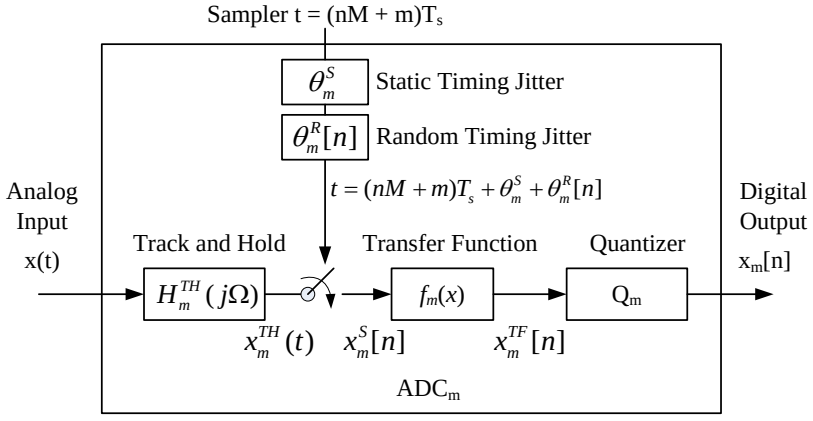

Figure 4: A practical ADC model.

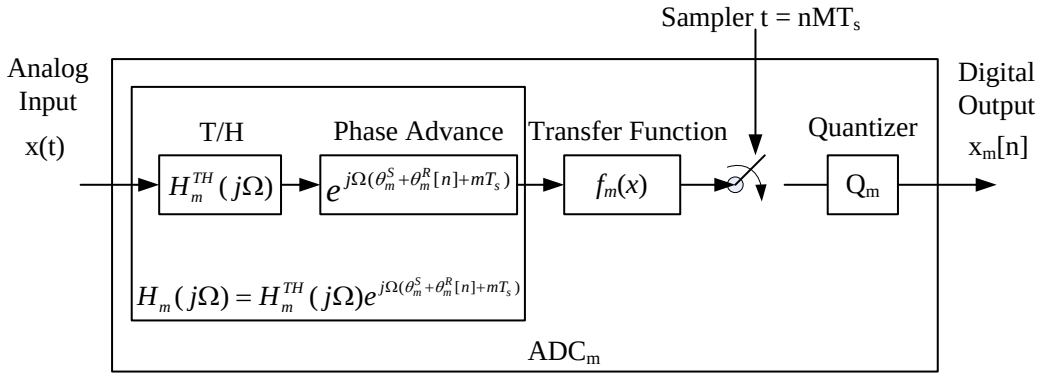

Figure 5: A re-organized ADC model.

In the digital end, the multiplexing operation of channel m can be decomposed into a factor M up-converter followed by a m sample delay element:

$$G_m(e^{j\omega}) = e^{-j\omega m}. \tag{21}$$

Finally, the output of each channel is added up to yield $y[n]$. The TIADC system can be re-structured as a NHFB structure as shown in Figure 6.

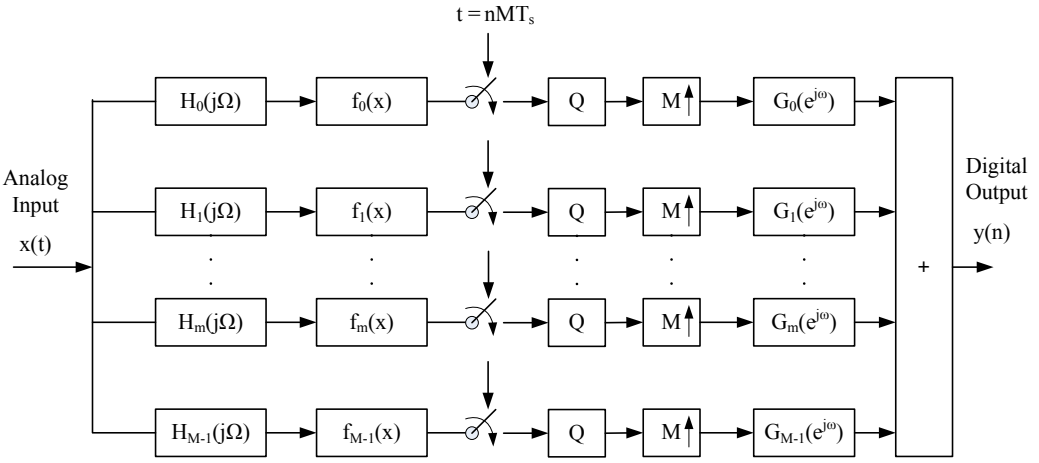

Figure 6: A non-linear hybrid filter bank structure.

When ignoring the quantization noise, denote $Y(e^{j\Omega T_s})$ as the discrete time Fourier transform (DTFT) of $y[n]$ and $X(j\Omega)$ as the continuous time Fourier transform (CTFT) of $x(t)$, the input-output relationship of the NHFB is given by [92].

$$Y(e^{j\Omega T_s}) = \frac{1}{T_s}\sum_{k=0}^{K-1}\sum_{p=-\infty}^{\infty} T_{p,k}(j\Omega), \tag{22}$$

where

$$T_{p,k}(j\Omega) = \frac{1}{M}\sum_{m=0}^{M-1} c_{k,m}(X(j(\Omega - p\frac{\Omega_s}{M})H_m(j\Omega - p\frac{\Omega_s}{M}))^{(*k)} G_m(e^{j\Omega T_s}) \tag{23}$$

and $\Omega_s = 2\pi F_s$. The notation $(*k)$ has the following frequency domain convolution definition. Suppose $Z(j\Omega)$ is the CTFT of $z(t)$, then $Z(j\Omega)^{(*k)}$ stands for the CTFT of $z(t)^k$ where

$$Z(j\Omega)^{(*k)} = \begin{cases} 2\pi\delta(\Omega) & k = 0 \\ Z(j\Omega) & k = 1 \\ Z(j\Omega) * Z(j\Omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} Z(j\lambda)Z(j(\Omega-\lambda))d\lambda & k = 2 \\ (\underbrace{Z * Z*, ..., *Z}_{\text{k-1 convolutions}})(j\Omega) & k > 2 \end{cases}. \tag{24}$$

According to (21), we can re-write (23) as

$$\begin{aligned} T_{p,k}(j\Omega) =& \frac{1}{M}\sum_{m=0}^{M-1} c_{k,m}(X(j(\Omega - p\frac{\Omega_s}{M})H_m(j\Omega - p\frac{\Omega_s}{M}))^{(*k)} e^{-jm\Omega T_s} \\ =& \frac{1}{M}\sum_{m=0}^{M-1} c_{k,m}(X(j(\Omega - p\frac{\Omega_s}{M})H_m(j\Omega - p\frac{\Omega_s}{M}))^{(*k)} e^{-j(\Omega - p\frac{\Omega_s}{M})mT_s} e^{-jpm\frac{2\pi}{M}} \\ =& \tilde{T}_{p,k}(j(\Omega - p\frac{\Omega_s}{M})), \end{aligned} \tag{25}$$

where

$$\tilde{T}_{p,k}(j\Omega) = \frac{1}{M}\sum_{m=0}^{M-1} c_{k,m}(X(j\Omega)H_m(j\Omega))^{(*k)} e^{-jm\Omega T_s} e^{-jpm\frac{2\pi}{M}}. \tag{26}$$

Equation (22) becomes

$$Y(e^{j\Omega T_s}) = \frac{1}{T_s}\sum_{k=0}^{K-1}\sum_{p=-\infty}^{\infty} \tilde{T}_{p,k}(j(\Omega - p\frac{\Omega_s}{M})). \tag{27}$$

$Y(e^{j\Omega T_s})$ is the superimposition of infinite terms of the shifted spectra of $\tilde{T}_{p,k}(j\Omega)$. To represent $Y(e^{j\Omega T_s})$ under the Nyquist range of $-\Omega_s/2 < \Omega < \Omega_s/2$, we need to

determine how many shifts of $\tilde{T}_{p,k}(j\Omega)$ reside inside the Nyquist range. Suppose the input signal $x(t)$ is band-limited according to

$$X(j\Omega) = 0 \quad |\Omega| \geq \Omega_s/2, \tag{28}$$

then $X(j\Omega)H_m(j\Omega)$ is also band-limited by $\Omega_s/2$. Its bandwidth expands by a time each time $X(j\Omega)H_m(j\Omega)$ convolves with itself.

$$X(j\Omega)H_m(j\Omega)^{(*k)} = 0 \quad |\Omega| \geq k\Omega_s/2. \tag{29}$$

To find the range of p that makes $\tilde{T}_{p,k}(j(\Omega - p\frac{\Omega_s}{M}))$ band-limited by $\Omega_s/2$, we can write

$$p_{\max}\frac{\Omega_s}{M} - k\frac{\Omega_s}{2} < \frac{\Omega_s}{2}, \tag{30}$$

$$p_{\max} = \lceil M(k+1)/2 \rceil - 1. \tag{31}$$

Equation (27) is re-written as

$$Y(e^{j\Omega T_s}) = \frac{1}{T_s}\sum_{k=0}^{K-1}\sum_{p=-p_{\max}}^{p_{\max}} \tilde{T}_{p,k}(j(\Omega - p\frac{\Omega_s}{M})), \quad |\Omega| < \Omega_s/2. \tag{32}$$

2.1.1 Mismatch Analysis on Simplified Models

It is difficult to understand the relationship between input spectrum and output spectrum through equation (32). This section gradually simplifies the models so that the spectrum distortions caused by the mismatches can be analyzed in a more direct manner.

The simplest model for a track and hold circuit is a RC circuit shown in Figure 7. The transfer function can be written as

$$H^{TH}(j\Omega) = \frac{1}{j\Omega\tau + 1}, \tag{33}$$

where $\tau = RC$. The amplitude and phase response of $H^{TH}(j\Omega)$ can be written as

$$H^{TH}(j\Omega) = A(\Omega)e^{j\phi(\Omega)}, \tag{34}$$

where

$$\begin{aligned} A(\Omega) =& \frac{1}{\sqrt{1+(\Omega\tau)^2}} \\ =& 1 - \frac{1}{2}(\Omega\tau)^2 + \frac{3}{8}(\Omega\tau)^4 - ... \\ =& 1 + A^{NL}(\Omega), \end{aligned} \tag{35}$$

$$\begin{aligned} \phi(\Omega) =& \tan^{-1}(\Omega\tau) \\ =& -(\Omega\tau) + \frac{1}{3}(\Omega\tau)^3 - \frac{1}{5}(\Omega\tau)^5 - ... \\ =& -(\Omega\tau) + \phi^{NL}(\Omega), \end{aligned} \tag{36}$$

$A^{NL}(\Omega)$ and $\phi^{NL}(\Omega)$ stand for the non-linear amplitude and phase component. The cutoff frequency of the circuit is determined by

$$\Omega_c = 1/\tau. \tag{37}$$

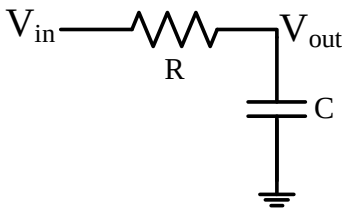

Figure 7: A model for the track and hold circuit.

The discrepancies of the time constant τ of the T/H circuits on each channel cause the bandwidth mismatches in the TIADC system. The analog analysis filter in (20) can be re-written as

$$H_m(j\Omega) = A_m(j\Omega)e^{-j(\tau_m - \theta_m^R[n] - \theta_m^S + mT_s)\Omega + \phi_m^{NL}(\Omega)}. \tag{38}$$

Denote the overall timing delay as $\Delta_{t,m} = \tau_m - \theta_m^R[n] - \theta_m^S$, (26) becomes

$$\tilde{T}_{p,k}(j\Omega) = \frac{1}{M}\sum_{m=0}^{M-1} c_{k,m}(X(j\Omega)A_m(\Omega)e^{j\phi_m^{NL}(\Omega)})^{(*k)}e^{-j\Delta_{t,m}\Omega}e^{-jpm\frac{2\pi}{M}}. \tag{39}$$

We can further simply the transfer function of the ADC on channel m as

$$f_m(x) = o_m + g_m x, \tag{40}$$

where o_m and g_m are the DC offset and gain of the ADC respectively. Then

$$\tilde{T}_{p,0}(j\Omega) = 2\pi\Lambda_{p,o}\delta(\Omega), \tag{41}$$

where

$$\Lambda_{p,o} = \frac{1}{M}\sum_{m=0}^{M-1} o_m e^{-jpm\frac{2\pi}{M}}. \tag{42}$$

$$\tilde{T}_{p,1}(j\Omega) = \Lambda_{p,bgt}(j\Omega)X(j\Omega), \tag{43}$$

where

$$\Lambda_{p,bgt}(j\Omega) = \frac{1}{M}\sum_{m=0}^{M-1} g_m A_m(\Omega) e^{j\phi_m^{NL}(\Omega)} e^{-j\Delta_{t,m}\Omega} e^{-jpm\frac{2\pi}{M}}. \tag{44}$$

The subscript o in $\Lambda_{p,o}$ indicates that this term is introduced by the DC offset mismatches. Similarly, the subscript bgt in $\Lambda_{p,bgt}$ indicates that this term is introduced by the bandwidth, gain and timing mismatches.

Equation (32) becomes

$$\begin{aligned} Y(e^{j\Omega T_s}) =& \frac{1}{T_s}\left[\sum_{p=-\lceil M/2\rceil+1}^{\lceil M/2\rceil-1} \Lambda_{p,o} 2\pi\delta(\Omega - p\frac{\Omega_s}{M})\right. \\ &\left. + \sum_{p=-M+1}^{M-1} \Lambda_{p,bgt}(j(\Omega - p\frac{\Omega_s}{M}))X(j(\Omega - p\frac{\Omega_s}{M}))\right], \quad |\Omega| < \Omega_s/2. \end{aligned} \tag{45}$$

Therefore, there will be $2\lceil M/2\rceil - 1$ impulses at $p\Omega_s/M$ as well as $2(M-1)$ aliasing terms in the output spectrum. Figure 8 and 9 shows the output spectra of a $M = 3$ and $M = 4$ TIADC system.

When the DC offset of all channels are equal: $o_m \equiv o$, then

$$\Lambda_{p,o} = o\sum_{n=-\infty}^{\infty}\delta[p - nM] \tag{46}$$

When the track and hold circuits, gains and timing delays in each channel are exactly matched: $A_m(\Omega) = A(\Omega), \phi_m^{NL} = \phi^{NL}, g_m = g, \Delta_{t,m} = \Delta$,

$$\Lambda_{p,bgt} = gA(\Omega)e^{j\phi^{NL}(\Omega)}e^{-j\Delta\Omega}\sum_{n=-\infty}^{\infty}\delta[p - nM]. \tag{47}$$

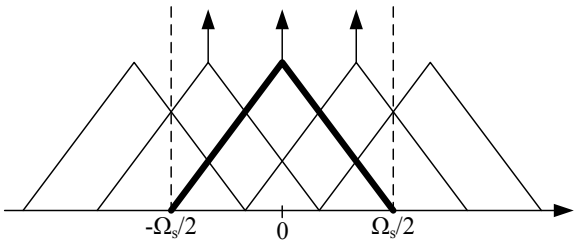

Figure 8: The output spectrum of a 3 channel TIADC system.

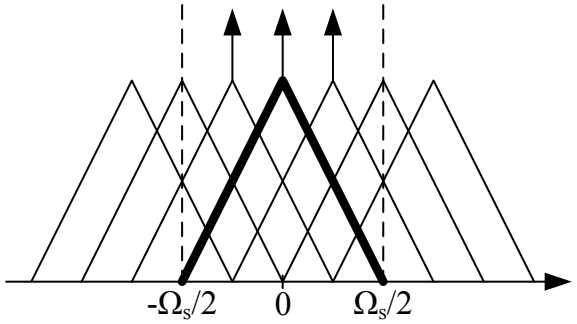

Figure 9: The output spectrum of a 4 channel TIADC system.

Equation (45) becomes

$$Y(e^{j\Omega T_s}) = \frac{1}{T_s}\left[o\pi\delta(\Omega) + gA(\Omega)e^{j\phi^{NL}(\Omega)}e^{-j\Delta\Omega}X(j\Omega)\right], \quad |\Omega| < \Omega_s/2. \tag{48}$$

Therefore, when the ADCs are exactly matched in a TIADC system, no aliasing terms will be present in the output spectrum over the Nyquist range.

Vogel [92] proposed a mismatch identification algorithm based on the simplified model in (45). Some special requirements were placed on the input signal spectrum. Figure 10 shows an example of the input spectrum to a 4 channel TIADC system. The input spectrum is required to contain spectrum holes at $p\frac{\Omega_s}{M}$ so that the original signal spectrum and the aliasing spectra won't be overlapped inside those holes. Also, it is also required that the input signal has spectral components close to $\pm\Omega_s/2$. These requirements allows us to separate and identify the envelop terms $\Lambda_{p,bgt}(j\Omega)$. The mismatch parameters can be evaluated based on $\Lambda_{p,bgt}(j\Omega)$.

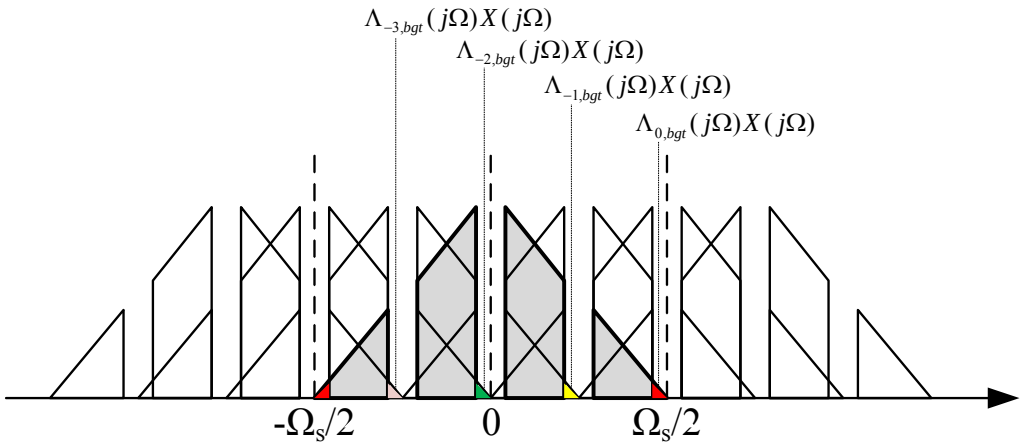

Figure 10: The output spectrum of a 4 channel TIADC system with a notched spectrum input.

2.2 Mismatch Calibrations for TIADC System

The calibration of mismatches in the time-interleaved system was carried out in two stages: identification and compensation. Since identification algorithms are not running permanently, their complexity and power consumption is not a critical concern. The compensation algorithms, on the other hand, are constantly running. Therefore, their efficiency is very important. The calibration can be executed either in the analog end [18] of the system or the digital end [24]. Methodologically, the calibration algorithms can be classified into active (or foreground) calibration and blind (or background) calibration. Under active calibration [61, 32, 33], a known pilot signal is sent to the TIADC system. The mismatches are identified according to the outputs of the system. Accurate and fast as active calibration is, it interrupts the normal converter operation, which is unacceptable in most real-time applications. Blind calibration method, on the contrary, performs calibration when the TIADC is in normal operation. Elbornsson proposed a blind timing mismatches identification algorithm based on a stationary assumption on the input signal in [19]. However, the estimate is biased which motivated him to include the method inside an adaptive system [21]. A comprehensive offset, gain and timing mismatch calibration algorithm was later

proposed in [20]. The disadvantage with Elbornsson's algorithm is that it usually need a huge amount of samples and thus has a high computational complexity.

Once the mismatches are identified, the next step is to compensate them. Gain and offset mismatches are relative easy to identify and compensate. Simple adders and multipliers are sufficient to serve this purpose. The compensation of timing mismatch is more complicated. Fractional delay filters [35, 57] can be applied to delay the samples by an arbitrary fractional amount of time so that the sampling timing can be matched perfectly in each channel. To avoid the re-design of fractional delay filters whenever the timing mismatches change, it is desirable to design an efficient filter structure to realize adjustable fractional delays.

This section proposes a coordinated blind adaptive calibration algorithm and a corresponding architecture to calibrate the TIADC channel mismatches. The proposed algorithm belongs to the blind algorithm category in a sense that only stationarity is assumed for the input signal. Its computational complexity and speed of convergence is shown to to be superior to the existing methods.

2.2.1 An Adaptive TIADC Mismatch Compensation Architecture

The estimation and adaptive compensation of DC offset, gain, and timing mismatches in TIADC systems can be integrated into one architecture as shown in Figure 11.

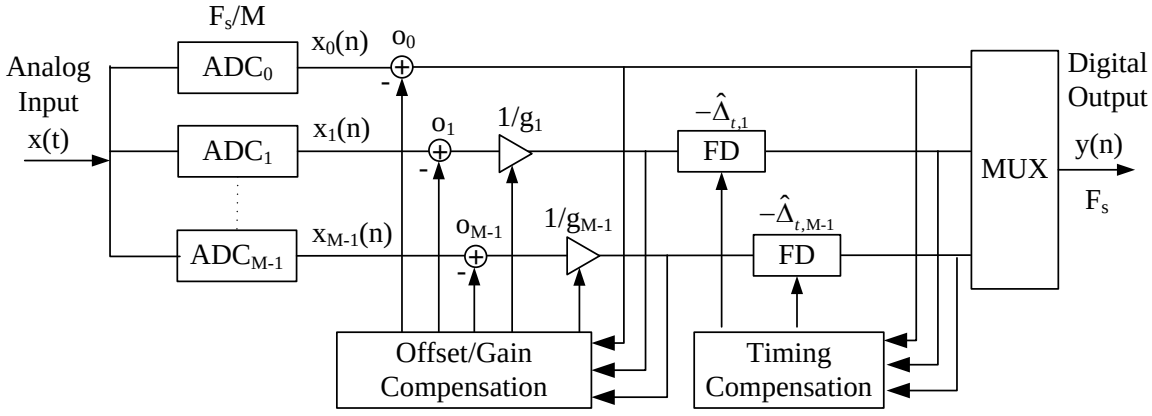

Figure 11: An adaptive digital TIADC calibration architecture.

In the proposed TIADC digital calibration architecture, channel 0 is set as the reference channel. DC offset and gain mismatches are relatively easy to estimate and compensate. They are corrected sequentially by the offset and gain compensation modules. Farrow-structured fractional delay (FD) filters are applied to make sure that the timing mismatches can be compensated in a continuous manner without an online filter design process. The timing mismatch compensation module requires a closed loop adaptation scheme. But we show that only one channel must be fully adapted. This simplification relies on a normalized relationship among the timing mismatches from different channels that can be established after an initialization step.

2.2.2 A Coordinated Blind Mismatch Calibration Algorithm

Suppose the input signal $x(t)$ is a zero mean, ergodic, and wide sense stationary process, the sampled signal at channel m is

$$x_m[n] = o_m + g_m x((nM+m)T_s + \Delta_{t,m}) \tag{49}$$
$$n = 0, ..., N-1,\ m = 0, ..., M-1,$$

where $o_m, g_m, \Delta_{t,m}$ are the DC offset, gain, and timing mismatches at channel m, N is the number of samples in each channel and M is the number of channels.

- DC offset mismatch estimation

The DC offset and gain mismatches are relatively easy to estimate and compensate by calculating the means and variances of the signals at each channel. The DC offset mismatches can be estimated according to

$$o_m = E[x_m[n]], \quad m = 0, ..., M-1. \tag{50}$$

- Gain mismatch estimation

After compensating the DC offset mismatches, the gain mismatches at each channel can be estimated according to

$$g_m = g_0 \sqrt{\frac{E[x_m^2[n]]}{E[x_0^2[n]]}}, \quad m = 1, ..., M-1, \tag{51}$$

where the gain g_0 in is assumed to be the gain in channel 0, which is set as the reference channel. In other words, we can only estimate the relative gains for channel 1 to $M-1$.

- Timing mismatch estimation

Once all the DC offsets are compensated and all the gains are calibrated relative to g_0, the most difficult task of timing mismatch estimation and compensation can be addressed. First, a cost function that equals the sum of squared cross-correlation differences between two consecutive pairs of adjacent channels is defined. When the timing mismatches are small compared to the sampling interval, an approximated normalization between the timing mismatches at each channel can be established by observing the cross-correlations of consecutive channels. Then, in the adaptation process, a steepest descent algorithm is executed to adjust the timing mismatch at a single adaptation channel, leaving the timing mismatches at the rest of the channels to be adjusted in a coordinated manner according to the normalized relationship established beforehand.

Note that

$$\begin{aligned} E[x_m[n]x_{m-1}[n]] =& g_0^2 E[x((nM+m)T_s + \Delta_{t,m})x((nM+m-1)T_s + \Delta_{t,m-1})] \\ =& g_0^2 R_x(T_s + \Delta_{t,m} - \Delta_{t,m-1}), \end{aligned} \tag{52}$$

where $R_x(\cdot)$ is the auto-correlation function of $x(t)$. When the timing mismatches are small, i.e. $|\Delta_{t,m}| \ll T_s$, we can approximate $R_x(T_s + \Delta_{t,m} - \Delta_{t,m-1})$ according to the first order Taylor expansion:

$$E[x_m[n]x_{m-1}[n]] \approx g_0^2 R_x(T_s) + g_0^2 R_x'(T_s)(\Delta_{t,m} - \Delta_{t,m-1}), \tag{53}$$

We can define a sub-cost function as

$$\begin{aligned} J_{m-1} =& E[x_m[n]x_{m-1}[n]] - E[x_{m-1}[n]x_{m-2}[n]] \\ \approx& g_0^2 R_x'(T_s)(\Delta_{t,m} - 2\Delta_{t,m-1} + \Delta_{t,m-2}), \\ & m = 2, ..., M-1, \end{aligned} \tag{54}$$

and we include the last channel by invoking wraparound

$$\begin{aligned} J_{M-1} =& E[x_0[n+1]x_{M-1}[n]] - E[x_{M-1}[n]x_{M-2}[n]] \\ \approx& g_0^2 R_x'(T_s)(\Delta_{t,0} - 2\Delta_{t,M-1} + \Delta_{t,M-2}). \end{aligned} \tag{55}$$

To unify the definition of (54) and (55), we can re-define the sub-cost function as

$$\begin{aligned} J_{m-1} = & E[x_{m \bmod M}[n + \left\lfloor \frac{m}{M} \right\rfloor] x_{m-1}[n]] - E[x_{m-1}[n]x_{m-2}[n]] \\ & m = 2, ..., M. \end{aligned} \tag{56}$$

Then we define the overall cost function as

$$J = \sum_{m=1}^{M-1} J_m^2. \tag{57}$$

An adaptive algorithm to update all channels can be derived from the gradient of the cost function. Suppose we are applying a delay $\hat{\Delta}_{t,m}$ to the signal $x_m[n]$ by means of digital fractional delay filters so that the new delay becomes $\Delta_{t,m} - \hat{\Delta}_{t,m}$. By denoting the compensated signal sequence as $\hat{x}_m[n]$ and the new sub and overall cost function after compensation as $\hat{J}_m$ and $\hat{J}$, we can evaluate the gradient of $\hat{J}$ according to (57, 54, 55):

$$\begin{aligned} \frac{\partial \hat{J}}{\partial \hat{\Delta}_{t,m}} =& -g_0^2 R_x'(T_s)(\hat{J}_{m+1} - 2\hat{J}_m + \hat{J}_{m-1}) \\ & m = 1, ..., M-1, \quad \hat{J}_0 = 0, \quad \hat{J}_M = 0. \end{aligned} \tag{58}$$

The exact value of $R'_x(T_s)$ is unknown. However, as long as we can determine the sign of $R'_x(T_s)$, we can still determine the direction of the gradient which is sufficient for the steepest descent algorithm to work.

Suppose $x(t)$ is a multi-sine signal with random phases:

$$x(t) = \sum_{p=1}^{P} a_p \sin(2\pi f_p t + \theta_p), \tag{59}$$

where P is the number of sines, $2f_p \leq F_s$, θ_p is independent from each other and uniformly distributed in $[0, 2\pi]$. The auto-correlation function of $x(t)$ is

$$R_x(\tau) = \sum_{p=1}^{P} \frac{a_p^2}{2} \cos(2\pi f_p \tau). \tag{60}$$

$$R'_x(T_s) = -\sum_{p=1}^{P} \frac{a_p^2}{2} \sin(2\pi f_p T_s) 2\pi f_p. \tag{61}$$

Therefore, $R'_x(T_s) \leq 0$ in this case. According to the steepest descent algorithm, we can update the estimate of $\hat{\Delta}_{t,m}$ by

$$\begin{aligned} \hat{\Delta}_{t,m}[k+1] &= \hat{\Delta}_{t,m}[k] - \mu_m(\hat{J}_{m+1}[k] - 2\hat{J}_m[k] + \hat{J}_{m-1}[k]) \\ & m = 1, ..., M-1, \end{aligned} \tag{62}$$

where k is the update step and μ_m is a positive step size. Therefore, for each update step k, we need to collect all the N samples of $x_m(n)$ for all M channels, evaluate $\hat{J}_m$, and update the delay estimates according to (62).

By assuming that the timing delays are constant during the adaptation process, we can significantly simplify the update scheme by the following coordinated adaptive algorithm. According to (54, 55), before compensation, we have

$$\begin{aligned} \gamma_1 = \frac{J_1}{J_2} &\approx \frac{\Delta_{t,2} - 2\Delta_{t,1}}{\Delta_{t,3} - 2\Delta_{t,2} + \Delta_{t,1}}, \\ \gamma_{m-1} = \frac{J_{m-1}}{J_m} &\approx \frac{\Delta_{t,m} - 2\Delta_{t,m-1} + \Delta_{t,m-2}}{\Delta_{t,m+1} - 2\Delta_{t,m} + \Delta_{t,m-1}}, \quad m = 3, .., M-2, \\ \gamma_{M-2} = \frac{J_{M-2}}{J_{M-1}} &\approx \frac{\Delta_{t,M-1} - 2\Delta_{t,M-2} + \Delta_{t,M-3}}{-2\Delta_{t,M-1} + \Delta_{t,M-2}}. \end{aligned} \tag{63}$$

Therefore, we have $M-2$ equations and $M-1$ variables $\Delta_{t,1}$ to $\Delta_{t,M-1}$. We cannot revolve the exact value of $\Delta_{t,1}$ to $\Delta_{t,M-1}$ due to the system being under-determined. However, we can normalize $\Delta_{t,m}$ with regard to a fixed adaptation channel. Without loss of generality, we can set channel 1 as our adaptation channel. Denote α_m as the normalization coefficients at channel m,

$$\Delta_{t,m} = \alpha_m \Delta_{t,1}, \quad m = 2, ..., M-1. \tag{64}$$

we can rewrite (63) as

$$\mathbf{\Gamma}\boldsymbol{\alpha} = -(\gamma_1 + 2)\mathbf{e}_1, \tag{65}$$

where,

$$\mathbf{\Gamma} = \begin{bmatrix} -(2\gamma_1+1) & \gamma_1 & & & & & \\ \gamma_2+2 & -(2\gamma_2+1) & \gamma_2 & & & & \\ & \ddots & \ddots & & & & \\ & -1 & \gamma_{m-1}+2 & -(2\gamma_{m-1}+1) & \gamma_{m-1} & & \\ & & \ddots & \ddots & \ddots & \ddots & \\ & & & -1 & \gamma_{M-2}+2 & -(2\gamma_{M-2}+1) \end{bmatrix}, \tag{66}$$

$$\boldsymbol{\alpha} = [\alpha_2, \alpha_3, ..., \alpha_{M-1}]^T, \tag{67}$$

$$\mathbf{e}_1 = [1, 0, ..., 0]^T. \tag{68}$$

After resolving $\boldsymbol{\alpha}$ during the initialization stage, we can coordinate the delay compensations as

$$\hat{\Delta}_{t,m} = \alpha_m \hat{\Delta}_{t,1} \quad m = 2, ..., M-1. \tag{69}$$

As a result, we just need to focus on adapting the time delay $\hat{\Delta}_{t,1}$ on channel 1, the delays for the other $M-1$ channels can be calculated directly from (69) instead of adaptively updated using (62). In other words, we only need to keep track of the data sequences $x_0[n]$ to $x_2[n]$. Not only does the proposed algorithm reduce computational

complexity, but it also accelerates the convergence speed of the adaptation. Imagining the overall cost function $\hat{J}$ as function of $\hat{\Delta}_{t,m}$ in a multi-dimensional space, the proposed algorithm will follow a straight line from the initial choice of $\hat{\Delta}_{t,m}$ to the optimality location of $\hat{J}$, which is the shortest path available.

Figure 12 shows two examples of TIADC timing mismatch calibration with the proposed coordinated adaptation algorithm. Every data block has $N = 5000$. The equivalent overall sampling frequency is $F_s = 1\,\text{kHz}$. The input multi-sine signal has frequencies at $[0.1055\ 0.0615\ 0.0505]F_s$. The timing mismatches are assumed to be uniformly distributed over $[-0.1\frac{T_s}{M}, 0.1\frac{T_s}{M}]$. The step size is set at $\mu_m = 5T_s$. Figure 12(a-c) shows an $M = 3$ case. The convergence path in Figure 12(a) follows a straight line from the initial choice of $\Delta_{t,1}$ and $\Delta_{t,2}$ towards the their true values due to the coordination of the proposed algorithm. Because of the shortened convergence path, the convergence speed of the cost function and estimated timing mismatches can be accelerated by a factor of 2.5 in Figure 12(b-c). Figure 12(d) shows a convergence speed comparison for $M = 8$ case.

2.3 Adjustable Digital Fractional Delay Filter Design

As the most critical building block in the proposed calibration architecture, the performance of the fractional delay filter directly impacts the quality of the timing mismatch compensations. This section concentrates on the design and implementation of digital fractional delay filters with an emphasis on the adjustability of the underlying structure, which requires that the filter structure be easy to update when changing the fractional delays.

2.3.1 Fractional Delay Filters

Digital fractional delay (FD) filters are widely used in many applications such as timing adjustment in digital modems [25], speech coding and synthesis [43], digital wave guide modeling [83] and sampling rate converters [34]. A comprehensive study

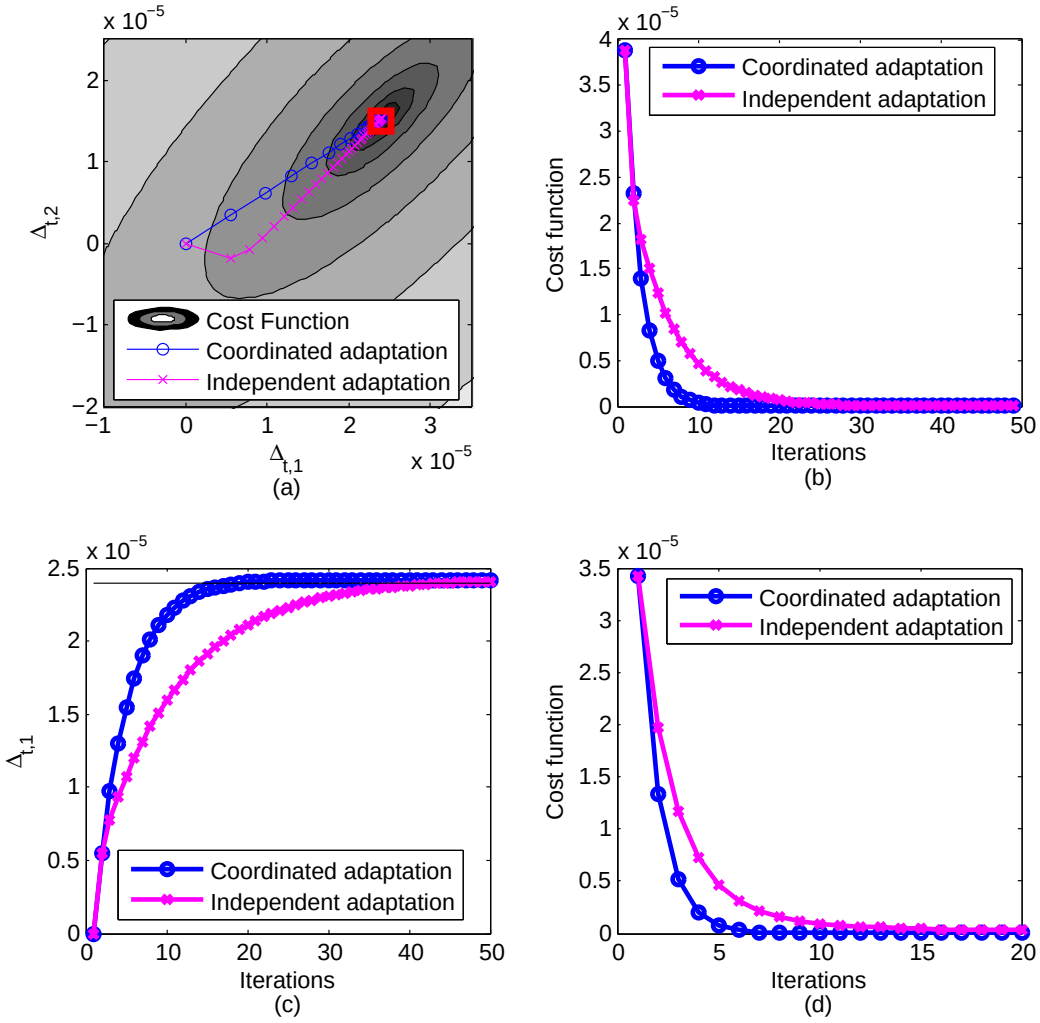

Figure 12: TIADC timing mismatch compensation examples. (a) convergence path over the proposed cost function ($M = 3$). (b) cost function convergence speed comparison ($M = 3$). (c) convergence speed comparison of timing mismatch estimation ($M = 3$). (d) cost function convergence speed comparison ($M = 8$).

of the existing design methodologies is summarized in [44, 66].

The compensation of timing mismatches in the proposed TIADC system can also be accomplished utilizing fractional delay filters. The transfer function of an ideal fractional delay filter can be written as

$$H_{\mathrm{id}}(z) = z^{-D}, \tag{70}$$

where D is the overall delay. The corresponding impulse response of the filter is

$$h_{\mathrm{id}}[n] = \mathrm{sinc}(n - D). \tag{71}$$

The impulse response of an ideal fractional delay filter is a shifted and sampled version of a sinc function. If D is an integer, the ideal impulse response is zero everywhere except at $n = D$. For non-integer values of D, the impulse response has infinite length. Figures 13 and 14 show the coefficients of two ideal FD filters with $D = 3$ and $D = 3.6$ respectively.

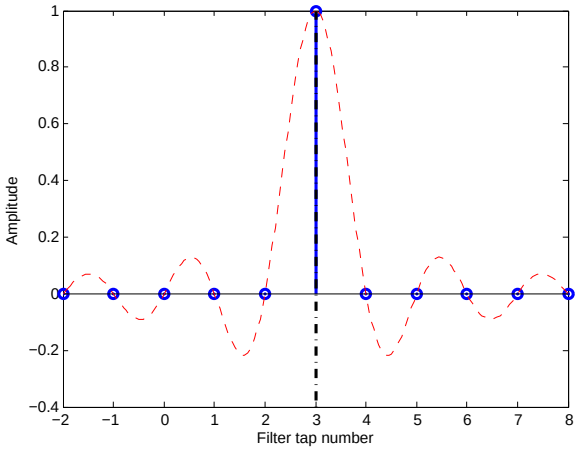

Figure 13: Ideal FD filter coefficients. D=3.

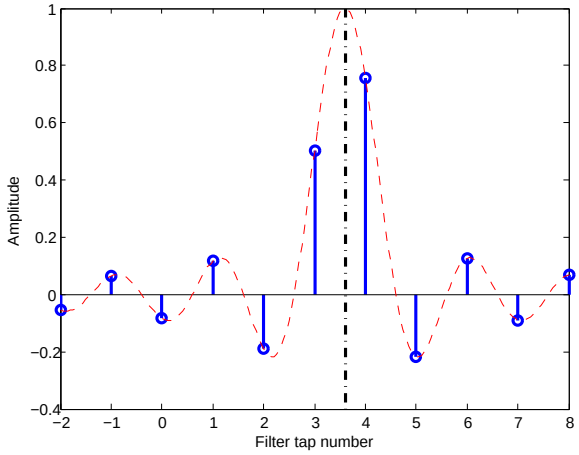

Figure 14: Ideal FD filter coefficients. D=3.6.

Various filter design criteria and their corresponding algorithms can be applied to the design of FD filters. For example, the minimum weighted least squares (WLS) error criterion, the maximally flat criterion and minimax criterion are typically adopted. A problem with the above mentioned design methods is that they are optimized for

a fixed value of D, which means that the filter has to be re-designed whenever D is changed. In TIADC systems, the timing mismatches among different channels might drift over time, which requires the compensation FD filters be flexible in adjusting its fractional delays online without a time consuming re-design process. The Farrow structure was invented to address this requirement.

2.3.2 Design Methods for Fractional Delay Filters

This section briefly reviews the design methods for FD filters with a fixed delay D.

2.3.2.1 Weighted Least Squares (WLS) Design Criterion

For an FIR filter of order N, the error function is defined as the difference between designed filter frequency response and ideal frequency response.

$$E(e^{j\omega}) = H(e^{j\omega}) - H_{\text{id}}(e^{j\omega}), \quad \omega \in [0, \alpha\pi], \quad 0 < \alpha \leq 1. \tag{72}$$

When α is closed to one, the delay will be matched over most of the frequency domain. The weighted least squares error is defined as

$$E_{LS} = \frac{1}{\pi} \int_0^{\alpha\pi} W(\omega) \left| E(e^{j\omega}) \right|^2 d\omega, \tag{73}$$

where $W(\omega)$ is a strictly positive weight function.

The design problem can be expressed in matrix vector notation if we define

$$\mathbf{h} = [\ h[0] \quad h[1] \quad \cdots \quad h[N]\ ]^T, \tag{74}$$

$$\mathbf{e} = [\ 1 \quad e^{-j\omega} \quad \ldots \quad e^{-j\omega N}\ ]^T, \tag{75}$$

$$\mathbf{E} = \Re\{\mathbf{e}\mathbf{e}^H\} = \begin{bmatrix} 1 & \cos(\omega) & \cdots & \cos(N\omega) \\ \cos(\omega) & 1 & & \cos((N-1)\omega) \\ \vdots & & \ddots & \vdots \\ \cos(N\omega) & \cos((N-1)\omega) & \cdots & 1 \end{bmatrix}. \tag{76}$$

Then (73) can be rewritten as

$$E_{LS} = \frac{1}{\pi}\int_0^{\alpha\pi} W(\omega)\left|H(e^{j\omega}) - H_{\mathrm{id}}(e^{j\omega})\right|^2 d\omega \tag{77}$$

$$= \frac{1}{\pi}\int_0^{\alpha\pi} W(\omega)\left[\mathbf{h}^T\mathbf{e} - H_{\mathrm{id}}(e^{j\omega})\right]\left[\mathbf{h}^T\mathbf{e} - H_{\mathrm{id}}(e^{j\omega})\right]^* d\omega \tag{78}$$

$$= \frac{1}{\pi}\int_0^{\alpha\pi} W(\omega)\left[\mathbf{h}^T\mathbf{E}\mathbf{h} - 2\mathbf{h}^T\Re\{H_{\mathrm{id}}(e^{j\omega})\mathbf{e}^*\} + \left|H_{\mathrm{id}}(e^{j\omega})\right|^2\right] d\omega \tag{79}$$

Next we define

$$\mathbf{A} = \frac{1}{\pi}\int_0^{\alpha\pi} W(\omega)\mathbf{E}d\omega, \tag{80}$$

$$\mathbf{b} = \frac{1}{\pi}\int_0^{\alpha\pi} W(\omega)\Re\{H_{id}(e^{j\omega})\mathbf{e}^*\}d\omega, \tag{81}$$

$$c = \frac{1}{\pi}\int_0^{\alpha\pi} W(\omega)\left|H_{id}(e^{j\omega})\right|^2 d\omega. \tag{82}$$

and then (73) becomes

$$E_{LS} = \mathbf{h}^T\mathbf{A}\mathbf{h} - 2\mathbf{h}^T\mathbf{b} + c \tag{83}$$

Therefore, the least squares error E_{LS} is minimized when

$$\mathbf{h} = \mathbf{A}^{-1}\mathbf{b}, \tag{84}$$

and the minimum value is

$$E_{LS\,\min} = c - \mathbf{b}^T\mathbf{A}^{-T}\mathbf{b}. \tag{85}$$

Note that the matrix $\mathbf{A}$ is independent from $H_{\mathrm{id}}(e^{j\omega})$, which means its inverse can be pre-computed. The change in $H_{\mathrm{id}}(e^{j\omega})$ only affects $\mathbf{b}$, and the filter coefficients $\mathbf{h}$ can be computed according to (84). If we define a special case where $\alpha = 1$, $W(\omega) \equiv 1$, $H_{\mathrm{id}}(e^{j\omega}) = e^{-jD\omega}$, then the elements of $\mathbf{A}$ and $\mathbf{b}$ become

$$\mathbf{A}_{ij} = \frac{1}{\pi}\int_0^{\pi} \cos((i-j)\omega)d\omega = \mathrm{sinc}(i-j) = \delta(i-j) \quad i,j = 0,...,N, \tag{86}$$

$$\mathbf{b}_i = \frac{1}{\pi}\int_0^{\pi} \cos((i-D)\omega)d\omega = \mathrm{sinc}(i-j) \quad i = 0,...,N. \tag{87}$$

In other words, **A** becomes an identity matrix and the filter coefficients will be

$$h[i] = \text{sinc}(i - D) \quad i = 0, ..., N, \tag{88}$$

$$E_{LS\,\text{min}} = 1 - \sum_{i=0}^{N} \text{sinc}^2(i - D). \tag{89}$$

Under this setup, the filter coefficients are a simply a truncated version of $h_{\text{id}}[n]$. The relationship between $E_{LS\,\text{min}}$ and the fractional delay D for an even-order filter ($N = 10$) is shown in Figure 15, and in 16 for an odd-order filter ($N = 11$). In both cases, the least squares error reaches its minimum when

$$(N - 1)/2 \leq D \leq (N + 1)/2. \tag{90}$$

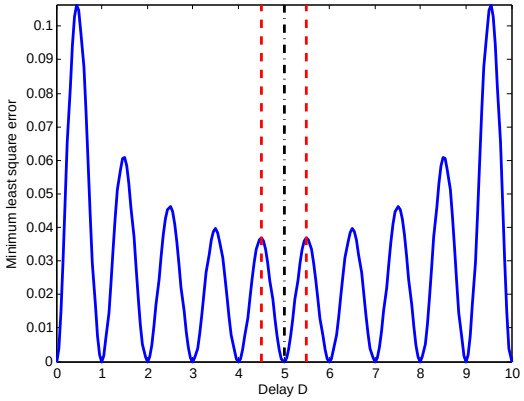

Figure 15: Minimum least squares error (N=10)

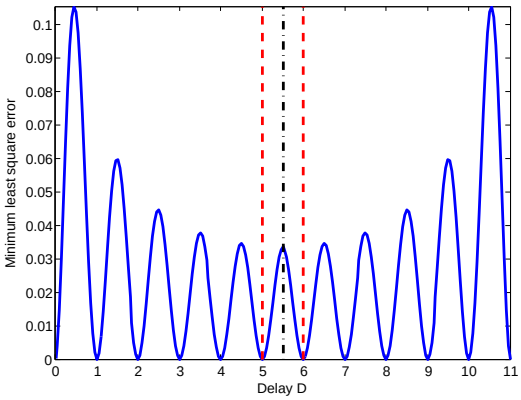

Figure 16: Minimum least squares error (N=11)

2.3.2.2 *Minimax Design Criterion*

Another design criterion, called the minimax design, is also widely used in the field of digital filter design. It aims to minimize the maximum approximation error. Instead of solving linear equations as in the WLS design case, minimax filter design problems are usually solved by iterative methods such as Remez exchange algorithm. [37] generalized the alternation theorem from real-only to complex case and proposed an efficient exchange algorithm for the design of complex valued and non-linear phase FIR filters in the Chebyshev case. A minimax FD filter design example is shown in Figures 17 through 19. For the case where delay $D = 13.8$ and the filter order is $N = 28$. The passband and stopband cutoff frequencies are $\omega_p = 0.8\pi$ and $\omega_s = 0.9\pi$, respectively. Both the magnitude resposne (Figure 18) and the group delay response (Figure 19) of the filter exhibit equal ripples deviating from the ideal frequency responses in the passband.

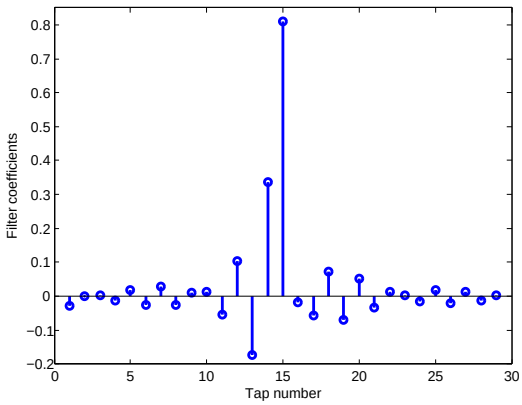

Figure 17: The filter coefficients of a FD filter designed with the minimax criterion. $D = 13.8, N = 28$.

2.3.2.3 *Maximally Flat Design Criterion*

For an FIR filter of order N, the maximally flat design criterion requires that the 0 to N-th order derivative of the error function as defined in (91) be zero at the frequency

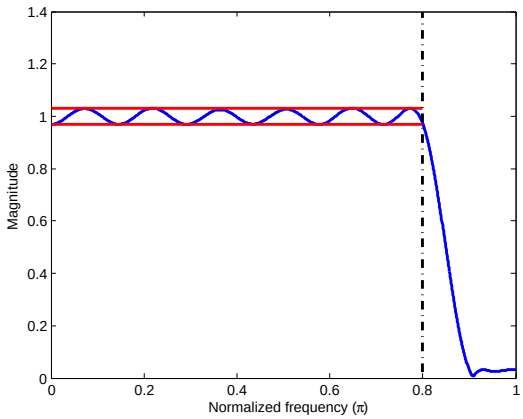

Figure 18: The magnitude response of a FD filter designed with the minimax criterion. $D = 13.8, N = 28$.

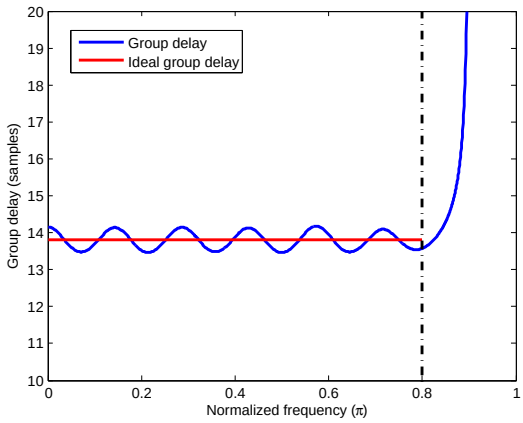

Figure 19: The phase response of a FD filter designed with the minimax criterion. $D = 13.8, N = 28$.

ω_0.

$$\left.\frac{d^k E(e^{j\omega})}{d\omega^k}\right|_{\omega=\omega_0} = 0, \quad k = 0, ..., N. \tag{91}$$

Accordingly,

$$\left.\frac{d^k}{d\omega^k}\left[\sum_{n=0}^{N} h(n)e^{-j\omega n} - e^{-j\omega D}\right]\right|_{\omega=\omega_0} = 0, \quad k = 0, ..., N. \tag{92}$$

$$\sum_{n=0}^{N} n^k h(n)e^{-j\omega_0 n} - D^k e^{-j\omega_0 D} = 0, \quad k = 0, ..., N. \tag{93}$$

If the error function is maximally flat at $\omega_0 = 0$, then (93) becomes

$$\sum_{n=0}^{N} n^k h(n) = D^k, \quad k = 0, ..., N. \tag{94}$$

Written in matrix form,

$$\mathbf{Nh} = \mathbf{D}, \tag{95}$$

where

$$\mathbf{N} = \begin{bmatrix} 0^0 & \cdots & N^0 \\ \vdots & \ddots & \vdots \\ 0^N & \cdots & N^N \end{bmatrix}, \tag{96}$$

$$\mathbf{D} = \begin{bmatrix} D^0 & \cdots & D^N \end{bmatrix}^T. \tag{97}$$

The matrix $\mathbf{N}$ is a Vandermonde matrix whose determinant is $\mathbf{N}$ is $N!(N-1)!$. Thus $\mathbf{N}$ is invertible and according to Cramer's rule, the solution can be obtained in the following explicit form:

$$h[n] = \prod_{\substack{k=0 \\ k \neq n}}^{N} \frac{D-k}{n-k}. \tag{98}$$

Equation (98) is also the Lagrange interpolation formula for equally spaced data points. Denote $h^D[n]$ as the FD filter coefficients with delay D, we have the following symmetry property:

$$h^D[n] = h^{N-D}[N-n], \quad n = 0, ..., N, \tag{99}$$

which means that a FD filter with delay $N - D$ is exactly the same as a FD filter with delay D with an inverse order of its filter coefficients, which implies that their magnitude responses are exactly identical.

2.3.3 The Farrow Structure

For a specific D, denote the order of the FD filter as N and coefficients as $h^D[k]$, Farrow [23] showed that the transfer function of $H(z)$ can be expressed as a polynomial in the delay D with coefficients $G_k(z)$ instead of a polynomial in z^{-1}.

$$H_{\mathrm{id}}(z) = z^{-D} = \sum_{k=0}^{N} G_k(z) D^k = \sum_{k=0}^{N} h^D[k] z^{-k}. \tag{100}$$

Equation (100) implies that the FD filter can be implemented in a structure shown in Figure 20. The input signal $x[n]$ is first passed through a bank of sub-filters $G_k(z)$. The outputs from those filters are cascaded through N multipliers by the constant D to yield the output $y[n]$. The sub-filter coefficients $g_k[n]$ are fixed after the filter design process. Thus delay through the structure is controlled by a single parameter D.

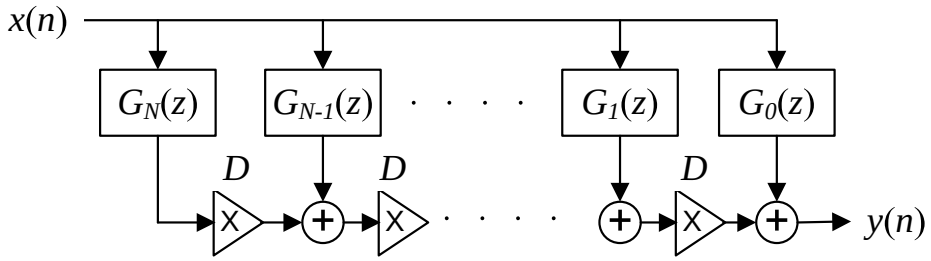

Figure 20: Block diagram of the original Farrow structure.

To obtain the sub-filters $G_k(z)$, the equality in (100) is required to hold for all integer values of D in the range of $D = 0, ..., N$. Therefore, we can write $N+1$ linear equations:

$$\mathbf{Vg} = \mathbf{z}, \tag{101}$$

where

$$\mathbf{V} = \begin{bmatrix} 0^0 & \cdots & 0^N \\ \vdots & \ddots & \vdots \\ N^0 & \cdots & N^N \end{bmatrix}, \tag{102}$$

$$\mathbf{g} = \begin{bmatrix} G_0(z) & \cdots & G_N(z) \end{bmatrix}^T, \tag{103}$$

$$\mathbf{z} = \begin{bmatrix} z^{-0} & \cdots & z^{-N} \end{bmatrix}^T. \tag{104}$$

Suppose each sub-filter $G_k(z)$ is an FIR filter of order N with coefficients $g_k[n]$:

$$G_k(z) = \sum_{n=0}^{N} g_k[n] z^{-n}, \tag{105}$$

then

$$\mathbf{g} = \mathbf{G}\mathbf{z}, \tag{106}$$

and (101) can be written as

$$\mathbf{V}\mathbf{G}\mathbf{z} = \mathbf{z}, \tag{107}$$

where the filter coefficients can be collected into a matrix

$$\mathbf{G} = \begin{bmatrix} g_0[0] & \cdots & g_0[N] \\ \vdots & \ddots & \vdots \\ g_N[0] & \cdots & g_N[N] \end{bmatrix}. \tag{108}$$

Therefore, the sub-filter coefficients matrix $\mathbf{G}$ can be obtained by inverting the Vandermonde matrix $\mathbf{V}$:

$$\mathbf{G} = \mathbf{V}^{-1}. \tag{109}$$

From (100), we also have

$$\mathbf{g}^T\mathbf{D} = \mathbf{z}^T\mathbf{h}, \tag{110}$$

$$\mathbf{h} = \mathbf{G}^T\mathbf{D}, \tag{111}$$

where

$$\mathbf{h} = \begin{bmatrix} h^D[0] & \cdots & h^D[N] \end{bmatrix}^T, \tag{112}$$

$$\mathbf{D} = \begin{bmatrix} D^0 & \cdots & D^N \end{bmatrix}^T. \tag{113}$$

Comparing equation (111) with (95), we can conclude that the original Farrow structured FD filter is exactly the same as the FD filter designed according to the maximally flat criterion. The optimal range of D lies within the 2 taps around the center of the impulse response.

$$\begin{aligned} N/2 - 1 \leq D \leq N/2 + 1, \quad \text{N even.} \\ N/2 - 0.5 \leq D \leq N/2 + 0.5, \quad \text{N odd.} \end{aligned} \tag{114}$$

There are two ways to decompose the overall delay D into an integer part plus a fractional delay.

$$D = \text{round}(D) + d, \quad -0.5 \leq d < 0.5. \tag{115}$$

$$D = \lfloor D \rfloor + d, \quad 0 \leq d < 1. \tag{116}$$

Logically, we can choose the filter order $N = 2\lfloor D \rfloor + 1$ or $N = 2\,\text{round}(D)$ to cover both the even and odd cases. To unify the expressions for the delay, we can write

$$D = N/2 + d, \quad -0.5 \leq d < 0.5. \tag{117}$$

In this case, equation (99) suggests that

$$h^{N/2+d}[n] = h^{N/2-d}[N - n]. \tag{118}$$

Therefore, we can conclude that

$$|H^{N/2+d}(e^{j\omega})| = |H^{N/2-d}(e^{j\omega})|, \tag{119}$$

$$\sphericalangle H^{N/2+d}(e^{j\omega}) = N - \sphericalangle H^{N/2-d}(e^{j\omega}), \tag{120}$$

where $\sphericalangle$ stands for the phase delay.

We can express $\mathbf{D}$ in terms of its fractional parts $\mathbf{d}$ as

$$\mathbf{D} = \mathbf{Q}\mathbf{d}, \tag{121}$$

where the elements of $\mathbf{Q}$ are given by

$$q_{ij} = \begin{cases} (N/2)^{j-i}\binom{j}{i} & 0 \leq i \leq j \leq N \\ 0 & 0 \leq j < i \leq N \end{cases}, \tag{122}$$

$$\mathbf{d} = \begin{bmatrix} d^0 & \cdots & d^N \end{bmatrix}^T. \tag{123}$$

It follows from (111) that

$$\mathbf{h} = \mathbf{G}^T\mathbf{Q}\mathbf{d} = \mathbf{G'}^T\mathbf{d}. \tag{124}$$

By pre-multiplying $\mathbf{G}$ by a matrix $\mathbf{Q}^T$, we obtain a new set of filter bank coefficients, namely $\mathbf{G}'$. Thereby, the delay parameter is changed from D to its fractional part d. The modified Farrow structure is shown in Figure 21.

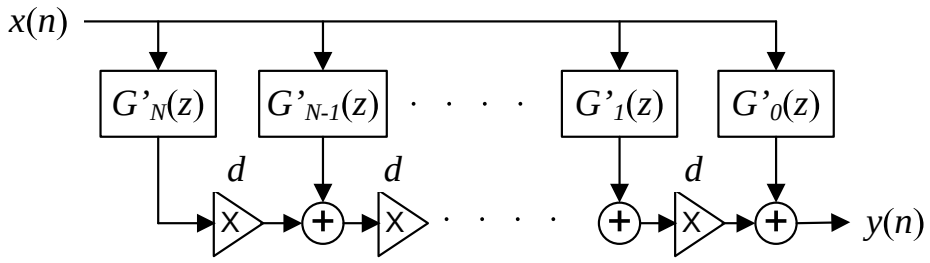

Figure 21: The block diagram of a modified Farrow structure.

The sub-filter coefficients and magnitude responses of $\mathbf{G}'$ for a Farrow structured FD filter of order $N = 5$ are shown in Figures 22 and 23. An even-order case $N = 6$ is shown in Figures 24 and 25. A general observation is that the sub-filters are either symmetric or anti-symmetric, both of which result in a linear phase response. The magnitude responses of the sub-filters $G'_k z$ have to approximate differentiators of ascending orders $\omega^k, k = 0, ...$

Figure 26 shows the magnitude responses of a Farrow structured FD filter with different fractional delays when the filter order is $N = 8$. The magnitude responses remain the same for d's and $-d$'s. So only positive d's are shown. Figure 27 shows the corresponding phase delay responses. The performance of the filter degrades as d approaches ± 0.5. Figure 27 shows the magnitude responses for a fixed fractional delay $d = 0.1$, but with different even orders. The corresponding phase delay responses are shown in Figure 29. Note that the $N/2$ phase delay part is removed for a fair comparison. The performance of the FD filter improves as the filter order N increases. Figures 30, 31, 32, 33 repeat the same experiments for odd filter orders.

Elegant as the structure is, there are some practical problems with the structure. First, the condition number of $\mathbf{V}$ increases exponentially as the order N increases, which means that it is difficult to design high order filters due to the numerical

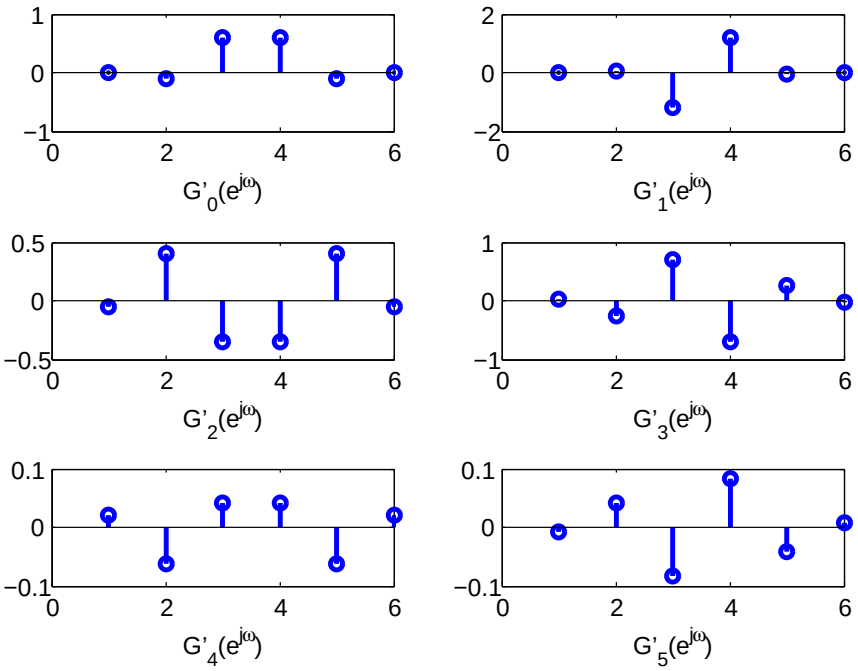

Figure 22: The sub-filter coefficients of $\mathbf{G}'$. $N = 5$.

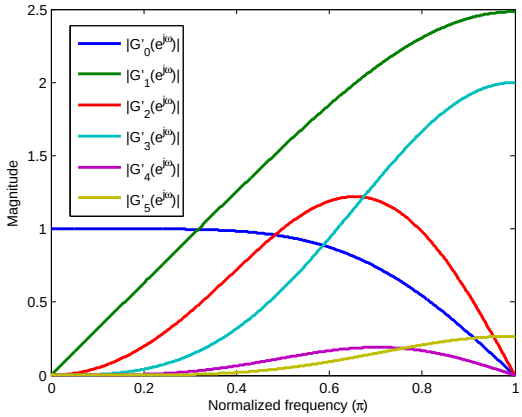

Figure 23: The magnitude responses of the sub-filters of $\mathbf{G}'$. $N = 5$.

instability of inverting $\mathbf{V}$. In MATLAB, we can only stably design the original Farrow structured FD filters for orders less than 30. Second, the cutoff frequency of the resulting FD filter depends on the filter order N and the fractional delay d. We can't simply extend the cutoff frequency by increasing the filter order due to the limit on the filter order. The approximation error of the original Farrow structured FD with regard to the ideal frequency response increases super-linearly as a function of d or ω.

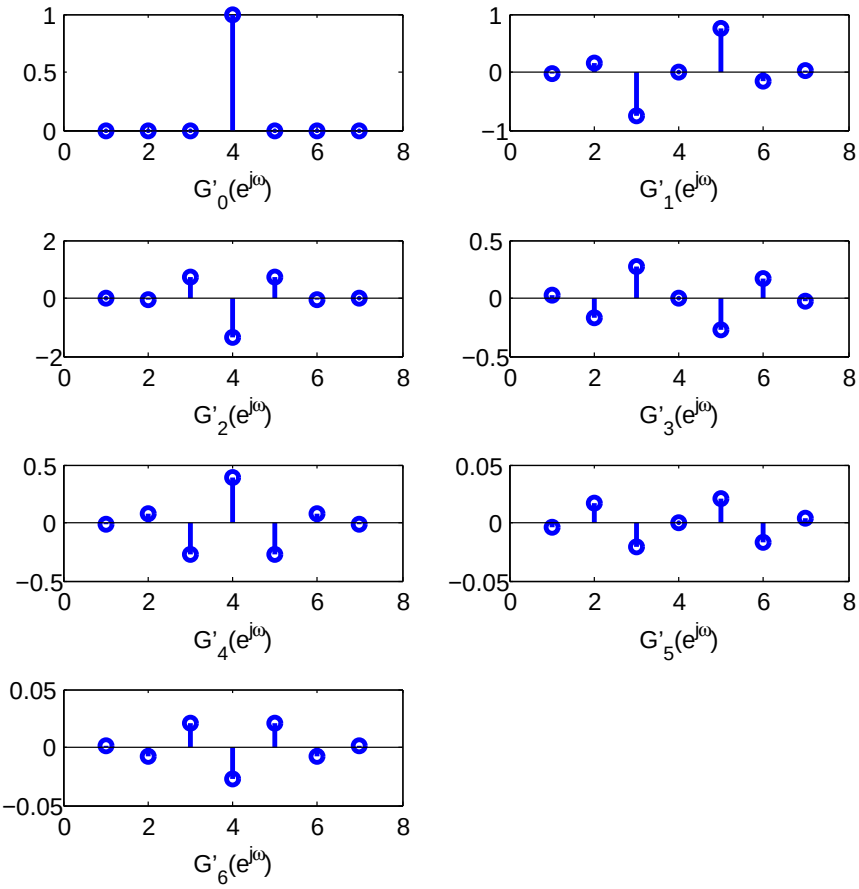

Figure 24: The sub-filter coefficients of $\mathbf{G}'$. $N = 6$.

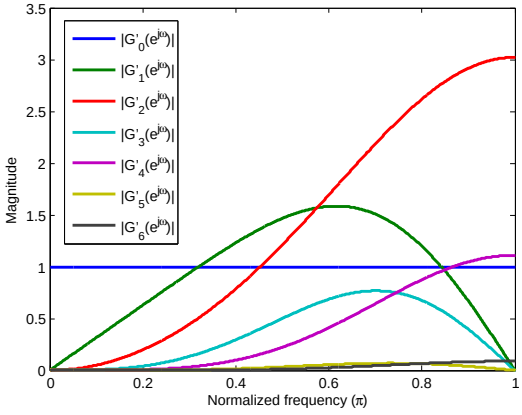

Figure 25: The magnitude responses of the sub-filters of $\mathbf{G}'$. $N = 6$.

The approximation errors of the original Farrow structured FD filters as a function of frequency ω with various combinations of N's and d's are shown in Figure 34. For a fixed d and N, the approximation error increases significantly under high frequency

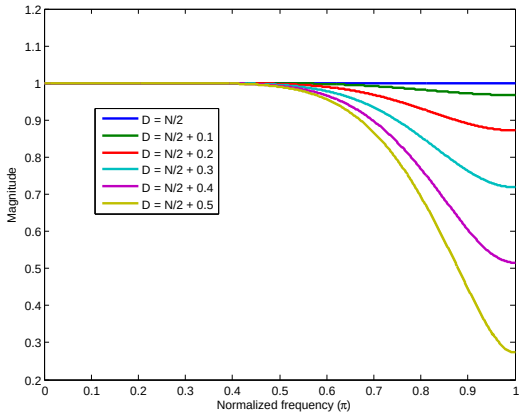

Figure 26: The magnitude responses of a Farrow structured FD filter with different fractional delays. $N = 8$.

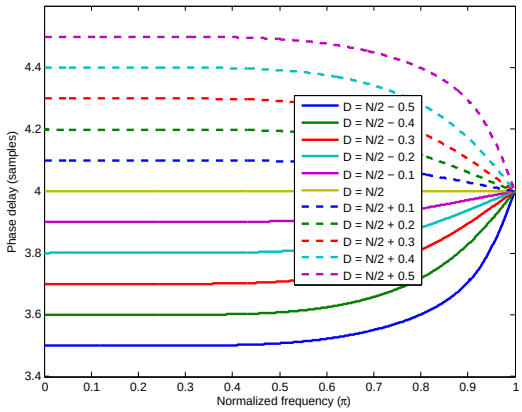

Figure 27: The phase delay responses of a Farrow structured FD filter with different fractional delays. $N = 8$.

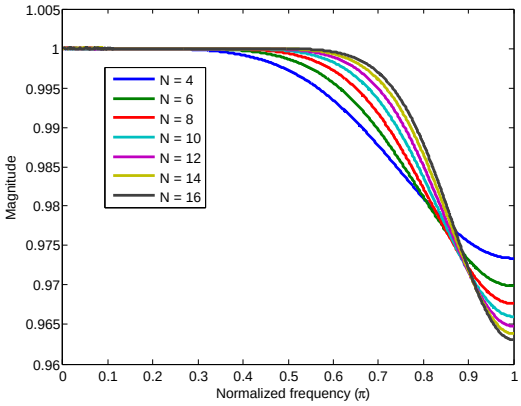

Figure 28: The magnitude responses of a Farrow structured FD filter with different even orders. $d = 0.1$.

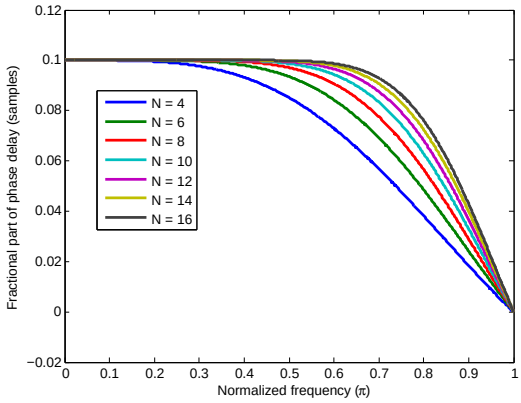

Figure 29: The fractional phase delay responses of a Farrow structured FD filter with different even orders. $d = 0.1$.

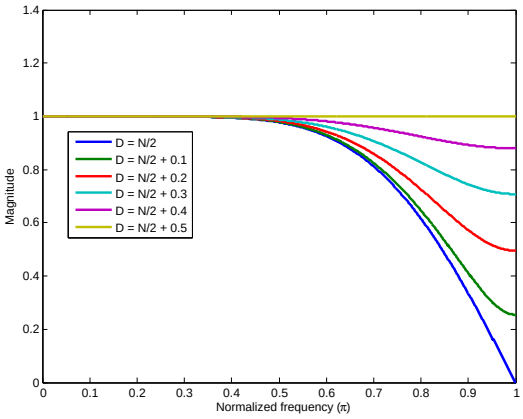

Figure 30: The magnitude responses of a Farrow structured FD filter with different fractional delays. $N = 7$.

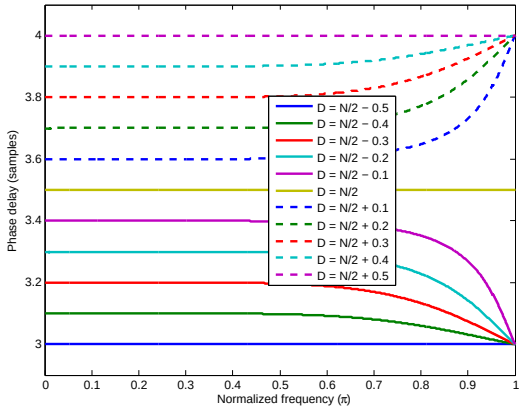

Figure 31: The phase delay responses of a Farrow structured FD filter with different fractional delays. $N = 7$.

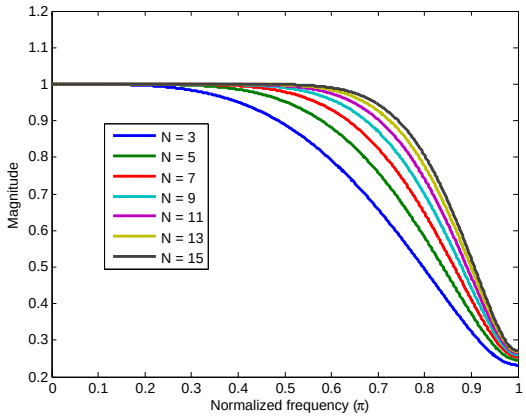

Figure 32: The magnitude responses of a Farrow structured FD filter with different odd orders. $d = 0.1$.

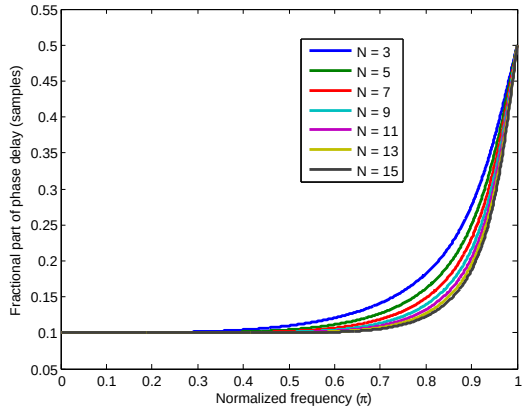

Figure 33: The fractional phase delay responses of a Farrow structured FD filter with different odd orders. $d = 0.1$.

bands. The maximum approximation errors over the frequency band $[0, 0.9\pi]$ of the original Farrow structured FD filter as a function of d with various N's are shown in Figure 35. For a fixed N and frequency band, the maximum approximation errors over the frequency band also increase dramatically as d increases.

Finally, the number of sub-filters always equals the sub-filter length in the original design method. However, we can see from the structure that the outputs from sub-filter $G'_k(z)$ is scaled by d^k. For large k or small d, the influence of the branch $G'_k(z)$ is negligible, which suggests the possibility of allowing less sub-filters.

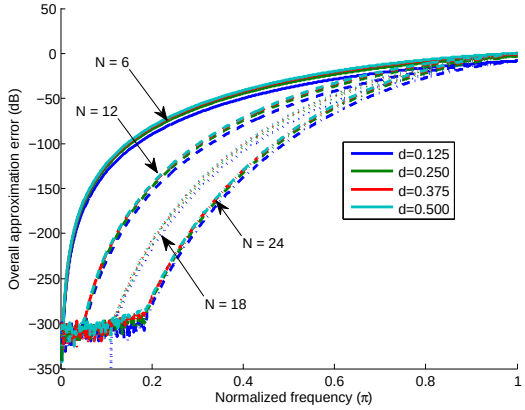

Figure 34: The approximation errors of the original Farrow structured FD filter as a function of ω.

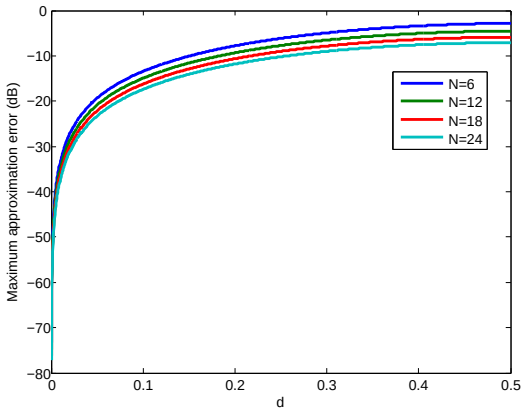

Figure 35: The maximum approximation errors over the frequency band $[0, 0.9\pi]$ of the original Farrow structured FD filter as a function of d.

2.3.4 Recursive Evaluation of the Filter Coefficient Matrix

The evaluation of the coefficient matrix $\mathbf{G}$ involves inverting a Vandermonde matrix $\mathbf{V}$ of dimension $N+1$, which has a computational complexity of $\mathcal{O}((N+1)^3)$. This section takes advantage of the special structure of $\mathbf{V}$ to derive a recursive formula to evaluate the filter coefficient matrix $\mathbf{G}$.

Consider the following general form of a Vandermonde matrix:

$$\mathbf{A} = \begin{bmatrix} 1 & x_0 & \cdots & x_0^N \\ 1 & x_1 & & x_1^N \\ \vdots & & \ddots & \vdots \\ 1 & x_N & \cdots & x_N^N \end{bmatrix}. \tag{125}$$

The matrix $\mathbf{V}$ in (102) is a special case of $\mathbf{A}$ where $x_i = i$. Vandermonde matrix is invertible as long as the elements x_i are distinct from each other. However, directly inverting the Vandermonde matrix $\mathbf{V}$ causes some numerical problems when N is very large. The matrix $\mathbf{A}$ can be factorized into upper and lower triangle factors $\mathbf{A} = \mathbf{LU}$. And $\mathbf{A}^{-1} = \mathbf{U}^{-1}\mathbf{L}^{-1}$. It has been derived in [77] that for this special case $x_i = i$,

$$\mathbf{U}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & -1 & 2 & \cdots \\ 0 & 0 & 1 & -3 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{126}$$

$$\mathbf{L}^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ -1 & 1 & 0 & 0 & \cdots \\ 1/2 & -1 & -1/2 & 0 & \cdots \\ -1/6 & 1/2 & -1/2 & 1/6 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}. \tag{127}$$

The general formula for the entries of the inverse of $\mathbf{U}^{-1}$ and $\mathbf{L}^{-1}$ can be written as

$$\mathbf{U}_{ij}^{-1} = \begin{cases} S(j,i) & 0 \le i \le j \le N \\ 0 & \text{Otherwise} \end{cases}, \tag{128}$$

$$\mathbf{L}_{ij}^{-1} = \begin{cases} \frac{(-1)^{i+j}}{j!(i-j)!} & 0 \le i \le j \le N \\ 0 & \text{Otherwise} \end{cases}, \tag{129}$$

where, $S(j,i)$ stands for the Stirling number of the first kind with the following definition:

$$x(x-1)...(x-n+1) = \sum_{k=0}^{n} S(n,k)x^k \quad 0 \le k \le n. \tag{130}$$

It's easy to prove that $S(n,k)$ has the following property:

$$S(n,0) = 0 \quad n \ge 1, \tag{131}$$

$$S(n,n) = 1 \quad n \ge 0, \tag{132}$$

$$S(n+1,k) = S(n,k-1) - nS(n,k). \tag{133}$$

Therefore, we can build up a table for S(n,k) using the the above recurrence relationship as shown in Table 1. Combining (128) and (129) we have

$$\mathbf{G}_{ij} = \sum_{k=0}^{N} \mathbf{U}_{ik}^{-1}\mathbf{L}_{kj}^{-1} = \begin{cases} \sum_{k=i}^{N} \mathbf{U}_{ik}^{-1}\mathbf{L}_{kj}^{-1} = \sum_{k=i}^{N} S(k,i)\frac{(-1)^{k+j}}{j!(k-j)!} & i > j \\ \sum_{k=i}^{N} \mathbf{U}_{ii}^{-1}\mathbf{L}_{ii}^{-1} = \sum_{k=i}^{N} S(k,i)\frac{(-1)^{k+i}}{i!(k-i)!} & i = j \\ \sum_{k=j}^{N} \mathbf{U}_{ik}^{-1}\mathbf{L}_{kj}^{-1} = \sum_{k=j}^{N} S(k,i)\frac{(-1)^{k+j}}{j!(k-j)!} & i < j \end{cases}, \tag{134}$$

which can be simplified to

$$\mathbf{G}_{ij} = \sum_{k=\max\{i,j\}}^{N} S(k,i)\frac{(-1)^{k+j}}{j!(k-j)!}. \tag{135}$$

Therefore, we can evaluate the filter coefficient matrix $\mathbf{G}$ explicitly.

For example, when $N = 3$,

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -11/6 & 3 & -3/2 & 1/3 \\ 1 & -5/2 & 2 & -1/2 \\ -1/6 & 1/2 & -1/2 & 1/6 \end{bmatrix}. \tag{136}$$

Table 1: Stirling number of the first kind

$n\backslash k$	0	1	2	3	4	$\cdots$
0	1					
1	0	1				
2	0	-1	1			
3	0	2	-3	1		
4	0	-6	11	-6	1	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

2.3.5 A Modified Farrow Structure

A key assumption during the development of the timing mismatch compensation algorithm is that $|\Delta_{t,m}| \ll T_s$, so the fractional delay $d = \frac{\Delta_{t,m}}{T_s}$ applied in the Farrow-structured adjustable FD filters is very small. However, the only design parameter in the original Farrow-structured FD filter is the filter order N. As a result, the original Farrow structure might be over-designed according to the analysis in Section 2.3.3. In this section, a new design strategy is presented to design the adjustable FD filters by optimizing over a given range of d and bandwidth ω_c.

According to (117), we can write the ideal frequency response of the fractional delay filter as

$$H_{\mathrm{id}}(e^{j\omega}) = e^{-jD\omega} = e^{-j(N/2)\omega} e^{-j\omega d}, \tag{137}$$

where ω is the normalized frequency. Given the passband of the resulting filter $\omega \in [0, \omega_c]$, where $\omega_c < \pi$ is the passband cutoff frequency, and the range of fractional delays $d \in [d_{\min}, d_{\max}]$, we can determine the range of the product $\omega d \in [(\omega d)_{\min}, (\omega d)_{\max}]$. Then we can approximate the fractional delay part as a complex polynomial of ωd, so that the approximation error is evenly distributed over the given range of ωd.

$$e^{-j\omega d} = \sum_{k=0}^{L} c_k (\omega d)^k + \varepsilon, \qquad |\varepsilon| \le \varepsilon_{\max}. \tag{138}$$

The complex coefficients c_k of the approximating polynomial will depend on the range of ωd. The optimal complex coefficients that minimize $\varepsilon_{\max}$ can be obtained

using a Remez exchange algorithm, or via convex optimization. Alternatively, we can approximate $e^{-j\omega d}$ by a Taylor expansion:

$$e^{-j\omega d} = \sum_{k=0}^{L} \frac{(-j)^k}{k!}(\omega d)^k + \varepsilon, \quad |\varepsilon| \leq \frac{|\omega d|^{L+1}}{(L+1)!}. \tag{139}$$

In this case, the worst-case approximation error can be estimated from L and ωd.

Figure 36 shows an example comparing the approximation errors of the optimal complex coefficient polynomial and the Taylor polynomial when the polynomial order is set to $L = 5$, the range of d is $[-0.2, 0.2]$, and the range of ω is $[0, 0.9\pi]$. The approximation error of the Taylor polynomial equals zero at $\omega d = 0$, but increases super-linearly as $|\omega d|$ increases. The approximation error of the optimal complex coefficient polynomial, on the other hand, is evenly distributed over the entire range of ωd. In terms of the maximum approximation error, the optimal complex coefficient polynomial is better than the Taylor polynomial when $|\omega d| > 0.1\pi$.

Table 2 shows the maximum approximation errors when $d \in [-d_{\max}, d_{\max}]$ and $\omega_c = 0.9\pi$. The maximum Taylor series approximation error increases as $d_{\max}$ increases. The maximum optimal polynomial approximation error is roughly $6L$ dB lower than the Taylor series approximation error at the same $d_{\max}$ and polynomial order L.

Table 3 shows the maximum optimal polynomial approximation errors when $d \in [0, d_{\max}]$. Since the maximum Taylor series approximation errors only depend on $\omega_c d_{\max}$, the values are the same as in Table 3. The reduced range of d enables the optimal polynomial to further reduce its maximum approximation error, which is roughly $12L + 6$ dB lower than the Taylor series approximation error.

According to (138), we can approximate $H_{\text{id}}(e^{j\omega})$ by $H_{\text{poly}}(e^{j\omega})$ such that

$$H_{\text{poly}}(e^{j\omega}) = e^{-j(N/2)\omega} \sum_{k=0}^{L} \left[(c_k^R \omega^k) d^k + (j c_k^I \omega^k) d^k \right], \tag{140}$$

$$|H_{\text{poly}}(e^{j\omega}) - H_{\text{id}}(e^{j\omega})| \leq \varepsilon_{\max}, \tag{141}$$

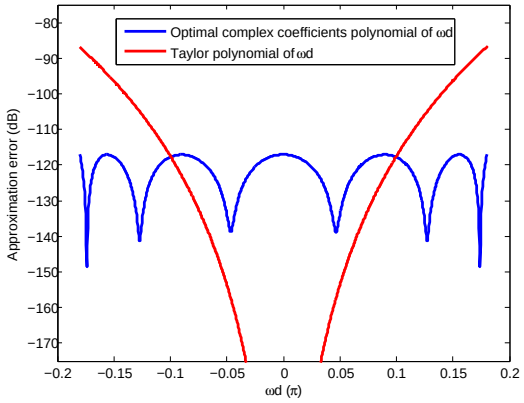

Figure 36: A comparison of the approximation errors to $e^{-j\omega d}$ when $L = 5$, $\omega_c = 0.9\pi$ and $d \in [-0.2, 0.2]$.

Table 2: Maximum approximation error (dB) to $e^{-j\omega d}$ for $d \in [-d_{\max}, d_{\max}]$ and $\omega_c = 0.9\pi$.

		$d_{\max}$			
		0.125	0.25	0.375	0.5
$L = 1$	Taylor	-24.12	-12.17	-5.28	-0.49
	Optimal	-30.20	-18.43	-11.85	-7.50
$L = 2$	Taylor	-42.69	-24.68	-14.22	-6.87
	Optimal	-54.77	-36.92	-26.69	-19.68
$L = 3$	Taylor	-63.75	-39.72	-25.70	-15.81
	Optimal	-81.86	-57.94	-44.12	-34.51
$L = 4$	Taylor	-86.76	-56.69	-39.14	-26.72
	Optimal	-110.88	-80.91	-63.53	-51.36
$L = 5$	Taylor	-111.36	-75.26	-54.17	-39.24
	Optimal	-141.34	-105.48	-84.55	-69.83
$L = 6$	Taylor	-137.29	-95.17	-70.55	-53.10
	Optimal	-173.22	-131.38	-106.92	-89.67

where, c_k^R and c_k^I stand for the real and imaginary parts of c_k. The error $\varepsilon_{\max}$ is referred to here as the first stage approximation error and it will be determined by the design parameter L. If we denote $G_k^R(e^{j\omega})$ and $G_k^I(e^{j\omega})$ as a pair of ideal linear phase k-th order differentiators with the following definitions:

$$G_k^R(e^{j\omega}) = e^{-j(N/2)\omega}(\omega^k), \tag{142}$$

$$G_k^I(e^{j\omega}) = je^{-j(N/2)\omega}(\omega^k), \tag{143}$$

Table 3: Maximum optimal polynomial approximation error (dB) to $e^{-j\omega d}$ when $d \in [0, d_{\max}]$ and $\omega_c = 0.9\pi$.

		$d_{\max}$			
		0.125	0.25	0.375	0.5
$L = 1$	Taylor	-24.12	-12.17	-5.28	-0.49
	Optimal	-42.17	-30.20	-23.27	-18.43
$L = 2$	Taylor	-42.69	-24.68	-14.22	-6.87
	Optimal	-72.78	-54.77	-44.29	-36.92
$L = 3$	Taylor	-63.75	-39.72	-25.70	-15.81
	Optimal	-105.90	-81.86	-67.84	-57.94
$L = 4$	Taylor	-86.76	-56.69	-39.14	-26.72
	Optimal	-140.63	-110.88	-93.33	-80.91
$L = 5$	Taylor	-111.36	-75.26	-54.17	-39.24
	Optimal	-176.24	-141.47	-120.41	-105.48
$L = 6$	Taylor	-137.29	-95.17	-70.55	-53.10
	Optimal	-206.58	-173.13	-148.40	-131.37

then $H_{\text{poly}}(e^{j\omega})$ can be written as

$$H_{\text{poly}}(e^{j\omega}) = \sum_{k=0}^{L} \left[c_k^R G_k^R(e^{j\omega}) + c_k^I G_k^I(e^{j\omega})\right] d^k. \tag{144}$$

Now, suppose we can design a pair of N^{th} order optimal linear phase FIR filters $\hat{G}_k^R(e^{j\omega})$ and $\hat{G}_k^I(e^{j\omega})$ that approximate $G_k^R(e^{j\omega})$ and $G_k^I(e^{j\omega})$ within the following maximum deviations over the frequency band $[0, \omega_c]$:

$$|\hat{G}_k^R(e^{j\omega}) - G_k^R(e^{j\omega})| \leq \delta_k^R, \tag{145}$$

$$|\hat{G}_k^I(e^{j\omega}) - G_k^I(e^{j\omega})| \leq \delta_k^I. \tag{146}$$

If we denote $\hat{H}(e^{j\omega})$ as the frequency response of the overall filter structure:

$$\hat{H}(e^{j\omega}) = \sum_{k=0}^{L} \left[c_k^R \hat{G}_k^R(e^{j\omega}) + c_k^I \hat{G}_k^I(e^{j\omega})\right] d^k, \tag{147}$$

then we can bound the overall approximation error by

$$|H_{\text{id}}(e^{j\omega}) - \hat{H}(e^{j\omega})| \leq \sum_{k=0}^{L} (|c_k^R|\delta_k^R + |c_k^I|\delta_k^I)|d|_{\max}^k + \varepsilon_{\max}. \tag{148}$$

The error term $\sum_{k=0}^{L} (|c_k^R|\delta_k^R + |c_k^I|\delta_k^I)|d|_{\max}^k$ is referred as the second stage approximation error and it will be determined by the design parameter N.

Table 4 summarizes the frequency response characteristics of the four types of linear phase FIR filters. We can set $\hat{G}_k^R(e^{j\omega})$ and $\hat{G}_k^I(e^{j\omega})$ as type I and III linear phase FIR filters, respectively, for even order N, or as type II and type IV filters for odd order N. Due to the restrictions at $\omega = \pi$, a full band approximation is unrealizable. Based on equation (147), a modified Farrow structure is shown in Figure 37. In the modified Farrow structure, the number of sub-filters $L+1$ determines the approximation error to $e^{-jd\omega}$. The sub-filter order N determines the approximation errors to the ideal differentiators. It is important to note that L and N are decoupled when using this new structure. Each sub-filter in the original Farrow structure is decomposed into two different types of linear phase filters $\hat{G}_k^R(e^{j\omega})$ and $\hat{G}_k^I(e^{j\omega})$ and synthesized with real coefficients c_k^R and c_k^I. Note that $\hat{G}_k^R(e^{j\omega}), \hat{G}_k^I(e^{j\omega}), c_k^R, c_k^I$ can all be designed offline for a given range of d and ω. The online adaptability of the structure is still controlled by the single parameter d.

According to (139), for the Taylor series approximation, the approximation coefficients are

$$c_k^R = \begin{cases} \frac{1}{k!} & k \bmod 4 = 0 \\ 0 & k \bmod 4 = 1 \\ \frac{-1}{k!} & k \bmod 4 = 2 \\ 0 & k \bmod 4 = 3 \end{cases} \qquad c_k^I = \begin{cases} 0 & k \bmod 4 = 0 \\ \frac{-1}{k!} & k \bmod 4 = 1 \\ 0 & k \bmod 4 = 2 \\ \frac{1}{k!} & k \bmod 4 = 3 \end{cases} \tag{149}$$

$$\varepsilon_{\max} = \frac{(\omega_c |d|_{\max})^{L+1}}{(L+1)!}. \tag{150}$$

For each k, either c_k^R or c_k^I is zero. Therefore, under the Taylor series approximation, only one sub-filter in each branch is active at a time.

2.3.5.1 Design Examples

The first example explores the behavior of the maximum deviations from the ideal frequency responses in a specified frequency band when the fractional delay d is

Table 4: The frequency responses of the four types of linear phase FIR filters.

	N	frequency response	restrictions
type I	even	$H(e^{j\omega}) = e^{-j\omega N/2}\sum_{k=0}^{N/2} a[k]\cos(\omega k)$ $a[0] = h[N/2]$ $a[k] = 2h[N/2-k] \quad k = 1, ..., N/2$	
type II	odd	$H(e^{j\omega}) = e^{-j\omega N/2}\sum_{k=0}^{(N+1)/2} b[k]\cos(\omega(k-1/2))$ $b[k] = 2h[(N+1)/2-k] \quad k = 1, ..., (N+1)/2$	$H(e^{j\pi}) = 0$
type III	even	$H(e^{j\omega}) = je^{-j\omega N/2}\sum_{k=0}^{N/2} c[k]\sin(\omega k)$ $c[k] = 2h[N/2-k] \quad k = 1, ..., N/2$	$H(e^{j0}) = 0$ $H(e^{j\pi}) = 0$
type IV	odd	$H(e^{j\omega}) = je^{-j\omega N/2}\sum_{k=0}^{(N+1)/2} d[k]\sin(\omega(k-1/2))$ $d[k] = 2h[(N+1)/2-k] \quad k = 1, ..., (N+1)/2$	$H(e^{j0}) = 0$

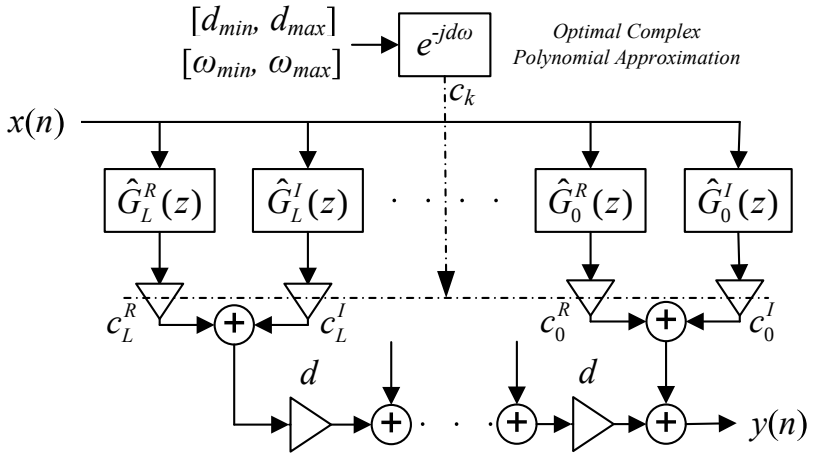

Figure 37: Block diagram of the modified Farrow structure.

varying. Consider a design specification where $\omega_c = 0.9\pi$, $d \in [0, 0.5]$. The maximum overall approximation errors over $[0, \omega_c]$ as a function of d with various combinations of L and N are shown in Figure 38. With the Taylor series approximation strategy, both the first stage and second stage approximation errors increase as d increases. As d gets larger, the approximation error is dominated by the first stage approximation error. On the other hand, with the optimal polynomial approximation strategy, the first stage approximation error is consistently bounded and significantly lower than its Taylor series approximation counterpart when d is large. Therefore, in terms of the maximum overall approximation error, the optimal polynomial approximation strategy is superior to the Taylor series approximation with the same specification of

N and L. As shown in Figure 38, the maximum approximation error is approximately the same for the optimal polynomial approximation technique with $N = 12$ and $L = 3$ and as the Taylor series approximation technique with $N = 18$ and $L = 4$. For the original Farrow structured FD filter, the maximum approximation error is much larger than the error from either Taylor or the optimal polynomial approximation strategy with the same filter order N.

In the second example, the overall approximation error in the frequency band $[0, \omega_c]$ for fixed values of d using the optimal polynomial and Taylor series approximation techniques are shown in Figures 39 and 40. Using the optimal polynomial approximation technique, the approximation errors exhibit an almost equal ripple behavior in $[0, \omega_c]$ when d is fixed. However, when using the Taylor series approximation technique, the approximation error increases significantly over the high frequency band when d gets larger. This is caused by the fact that a large value of the ωd product makes the first stage approximation error dominate the overall approximation error. For the original Farrow structured FD filter, the approximation errors are much lower than the Taylor and optimal polynomial approximation strategies over the low frequency band but rises considerably over the high frequency band. In terms of the maximum overall approximation error over the frequency band $[0, \omega_c]$, the optimal polynomial approximation strategy performs the best. Finally, it is worthwhile to point out that even better performance can be achieved for specifications such as a bandpass frequency band or an asymmetric range of d.

Consider another design specification where the maximum overall approximation error e of the FD filter is given. The objective is to find the optimal combination of N and L that satisfies the error constant e. Note from equation (148) that for each branch k, the differentiator approximation errors δ_k^R and δ_k^I are scaled by $|c_k^R||d|_{\max}^k$ and $|c_k^I||d|_{\max}^k$, respectively, which implies that we can relax δ_k^R and δ_k^I when $|c_k^R||d|_{\max}^k$ and $|c_k^I||d|_{\max}^k$ are small. In other words, the differentiators don't have to have the

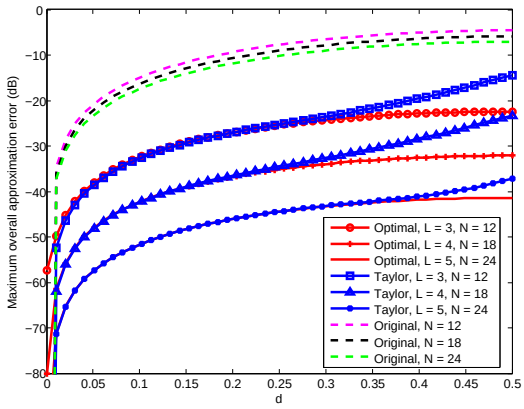

Figure 38: Maximum overall approximation error comparison as a function of d for $d \in [0, 0.5]$ and $\omega_c = 0.9\pi$.

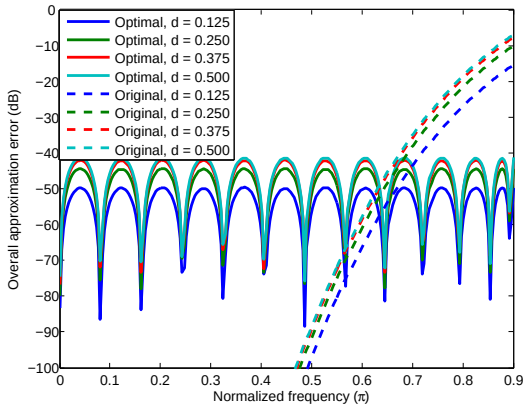

Figure 39: Overall approximation error over $[0, \omega_c]$ using optimal polynomial approximation when $d \in [0, 0.5]$, $\omega_c = 0.9\pi$, $N = 24$ and $L = 4$.

same length. We propose a design strategy as summarized in Algorithm 1 to find the optimal differentiator orders so that the scaled differentiator approximation errors in each branch are equal. As a special case, for even-order filters, $\hat{G}_0^R(e^{j\omega})$ is supposed to approximate $G_0^R(e^{j\omega}) = e^{-j(N/2)\omega}$ where the approximation error δ_0^R is exactly zero. We can attribute the differentiator approximation error in channel 0 solely to the $G_0^I(e^{j\omega})$ sub-branch by setting $\delta_0^I = \frac{2\bar{\delta}}{|c_k^R|}$. For the Taylor polynomial approximation strategy, the synthesizing coefficients are fixed, so the design algorithm is modified as shown in Algorithm 2. Similarly, we have $\delta_0 = 0$ for even-order case. Accordingly, we can modify equation (155) to $\bar{\delta} = \frac{e - \varepsilon_{\max}}{L}$.

For an example where the filter design specifications are $e = 0.001$, $d \in [0, 0.5]$ and

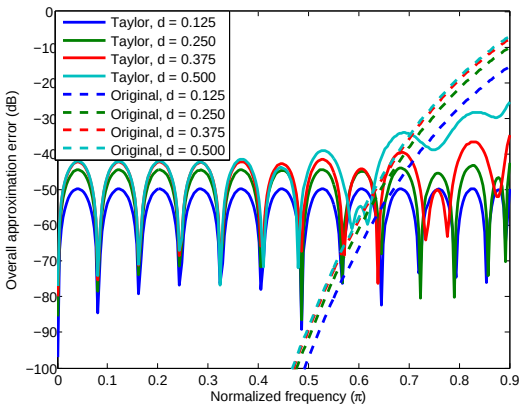

Figure 40: Overall approximation error over $[0, \omega_c]$ using Taylor series approximation when $d \in [0, 0.5]$, $\omega_c = 0.9\pi$, $N = 24$ and $L = 4$.

$\omega_c = 0.9\pi$. The resulting filter orders and maximum deviations using the optimal and Taylor polynomial approximation strategies are summarized in Tables 5 and 6. The deviations δ and $\hat{\delta}$ denote the prescribed and designed deviations of the differentiators. Since this example uses filter designs with even order filters, N_0^R and N_0 can be any even integer so they are crossed out in the tables. The corresponding overall approximation error in the frequency band $[0, \omega_c]$ for fixed values of d are shown in Figures 41 and 42. The maximum approximation errors of both approximation strategies are 10 dB more conservative than the prescribed -60 dB error bound. The optimal polynomial approximation strategy uses only five sub-filters with a maximum sub-filter order of 54, while the Taylor polynomial approximation strategy uses nine sub-filters with a maximum sub-filter order of 56. Therefore, the optimal polynomial approximation strategy is more economic than the Taylor polynomial case.

Table 5: Design example with the modified Farrow structure using the optimal polynomial approximation strategy.

k	δ_k^R	$\hat{\delta}_k^R$	N_k^R	δ_k^I	$\hat{\delta}_k^I$	N_k^I
0	0	0	X	2.659	1.000	2
1	9.365e-2	8.609e-2	6	1.825e-4	1.634e-4	54
2	7.134e-4	5.437e-4	30	2.474e-2	2.081e-2	32
3	4.067e-2	3.310e-2	18	3.693e-3	3.128e-3	50
4	4.732e-2	3.682e-2	24	5.540e-2	5.236e-2	40

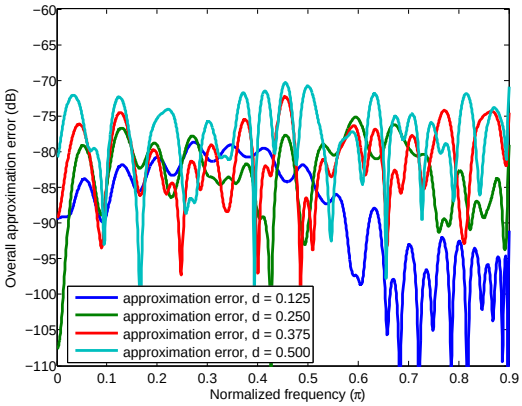

Figure 41: Overall approximation error in $[0, \omega_c]$ using the optimal polynomial approximation when $d \in [0, 0.5]$ and $\omega_c = 0.9\pi$.

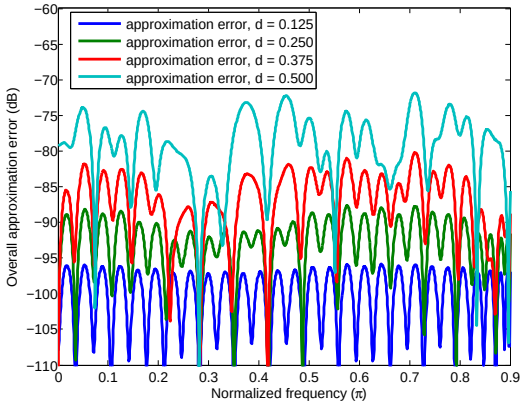

Figure 42: Overall approximation error in $[0, \omega_c]$ using the Taylor series approximation when $d \in [0, 0.5]$ and $\omega_c = 0.9\pi$.

2.3.6 Scalable Multirate Filter Based on the Farrow Structure

The power of the Farrow structure is not limited to the design of adjustable fractional delay filters. Its single parameter adjustability can also be applied to the design of scalable multirate filters [34]. Consider the interpolation by M structure as shown in Figure 43. The input signal is processed by a factor of M up-converter, followed by an FIR lowpass filter $H(z)$ with cutoff frequency at π/M. If we denote the order of $H(z)$ as N, $H(z)$ should approximate z^{-N} in the frequency band $\omega \in [0, \pi/M]$. With a polyphase decomposition, we can process $x[n]$ by M parallel filters $H_m(z)$ whose

Table 6: Design example with the modified Farrow structure using the Taylor polynomial approximation strategy.

k	δ_k	$\hat{\delta}_k$	N_k
0	0	0	X
1	1.511e-4	1.170e-4	56
2	6.043e-4	5.437e-4	30
3	3.626e-3	3.128e-3	50
4	2.901e-2	2.419e-2	26
5	2.901e-1	2.291e-1	38
6	3.481	3.116	16
7	4.873e+1	4.857e+1	20
8	7.797e+2	4.776e+2	6

outputs are interleaved. The relationship between $H(z)$ and $H_m(z)$ can be written as

$$H(z) = \sum_{m=0}^{M-1} z^{-m} H_m(z^M), \tag{158}$$

which implies that each component $z^{-m}H_m(z^M)$ in the summation should approximate approximate $\frac{1}{M}z^{-N}$ in the frequency band of $[0, \pi/M]$. We will have the following flow of logic:

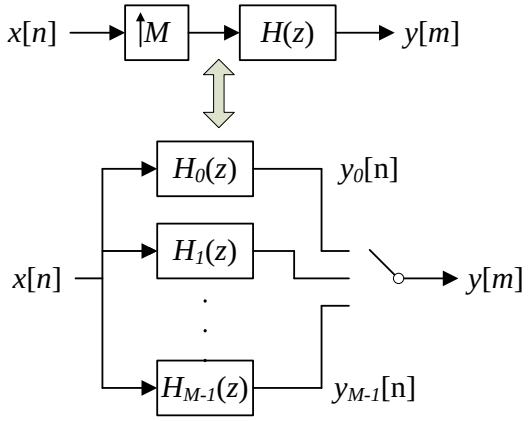

Figure 43: The polyphase decomposition of a factor M interpolator.

$$z^{-m}H_m(z^M) \to \frac{1}{M}z^{-N/2}, \quad \omega \in [0, \pi/M] \tag{159}$$

$$H_m(z^M) \to \frac{1}{M}z^{-N/2+m}, \quad \omega \in [0, \pi/M] \tag{160}$$

$$H_m(z) \to \frac{1}{M}z^{-(N/2-m)/M}, \quad \omega \in [0, \pi] \tag{161}$$

Input: maximum overall approximation error e, cutoff frequency ω_c, fractional delay range $[d_{\min}, d_{\max}]$.
Output: sub-filter coefficients $\hat{H}_k(e^{j\omega})$.

Determine the number of sub-filters:
Find the smallest optimal polynomial order L such that $\varepsilon_{\max} < e$. Calculate the average scaled differentiator approximation error

$$\bar{\delta} = \frac{1}{2}\frac{e - \varepsilon_{\max}}{L+1}. \tag{151}$$

Determine the orders of the sub-filters:
Calculate the differentiator approximation errors:

$$\delta_k^R = \frac{\bar{\delta}}{|c_k^R||d|_{\max}^k}, \tag{152}$$

$$\delta_k^I = \frac{\bar{\delta}}{|c_k^I||d|_{\max}^k}. \tag{153}$$

Design $\hat{G}_k^R(e^{j\omega})$ and $\hat{G}_k^I(e^{j\omega})$ with minimum orders N_k^R and N_k^I to approximate $G_k^R(e^{j\omega}) = e^{-j(N_k^R/2)\omega}\omega^k$ and $G_k^I(e^{j\omega}) = je^{-j(N_k^I/2)\omega}\omega^k$ whose approximation errors are bounded by δ_k^R and δ_k^I.
Each sub-filter $\hat{H}_k(e^{j\omega})$ can be synthesized by

$$\hat{H}_k(e^{j\omega}) = c_k^R\hat{G}_k^R(e^{j\omega}) + c_k^I\hat{G}_k^I(e^{j\omega}). \tag{154}$$

Algorithm 1: Design algorithm for the modified Farrow structured FD filter with optimal polynomial approximation strategy.

Ideally, each sub-filter $H_m(z)$ is a fractional delay filter with delay $D_m = (N/2 - m)/M$. If we denote the filter order of $H_m(z)$ as N_s and set $N_s = \frac{N-M}{M}$, then

$$D_m = N_s/2 + d_m, \tag{162}$$

where

$$d_m = \frac{1}{2} - \frac{m}{M}. \tag{163}$$

The value of d_m will lie in the range $[-\frac{1}{2} + \frac{1}{M}, \frac{1}{2}]$ and will be anti-symmetric:

$$d_m = -d_{M-m}. \tag{164}$$

Therefore, $H_m(z)$ can share the same sub-filters in the Farrow structure. The only difference lies in the fractional delay d_m. Denote $G_k(z), k = 0, ..., L$ as the sub-filters

Input: maximum overall approximation error e, cutoff frequency ω_c, fractional delay range $[d_{\min}, d_{\max}]$.
Output: sub-filter coefficients $\hat{H}_k(e^{j\omega})$

Determine the number of sub-filters:
Find the smallest Taylor polynomial order L such that $\frac{(\omega_c|d|_{\max})^{L+1}}{(L+1)!} < e$.
Calculate the average scaled differentiator approximation error

$$\bar{\delta} = \frac{e - \varepsilon_{\max}}{L+1}. \tag{155}$$

Determine the orders of the sub-filters:
Calculate the differentiator approximation errors:

$$\delta_k = \frac{\bar{\delta} k!}{|d|^k_{\max}}, \tag{156}$$

Design $\hat{G}_k(e^{j\omega})$ with minimum orders N_k to approximate
$G_k(e^{j\omega}) = (-j)^k e^{-j(N_k/2)\omega}\omega^k$ whose approximation errors are bounded by δ_k.
Each sub-filter $\hat{H}_k(e^{j\omega})$ is given by by

$$\hat{H}_k(e^{j\omega}) = \frac{1}{k!}\hat{G}_k(e^{j\omega}) \tag{157}$$

Algorithm 2: Design algorithm for the modified Farrow structured FD filter with Taylor polynomial approximation strategy.

in the Farrow structure, we can write $H_m(z)$ as

$$H_m(z) = \sum_{k=0}^{L} G_k(z) d_m^k. \tag{165}$$

Taking advantage of the symmetry of d_m in equation (164), we have

$$H_{M-m}(z) = \sum_{k=0}^{L} G_k(z)(-d_m)^k. \tag{166}$$

As a result, we can implement $H_m(z)$ and $H_{M-m}(z)$ simultaneously as shown in Figure 44. Denote the filter bank $G_k(z)$ as $\mathbf{G}(z)$ and the multiplication cascade as $\mathbf{d}_m^M$ in Figure 44. The complete interpolator structure is shown in Figure 45. The filter bank $\mathbf{G}(z)$ is shared by all $H_m(z)$. The outputs of $\mathbf{G}(z)$ are processed by $\lfloor M/2 \rfloor$ multiplication cascades $\mathbf{d}_m^M$, which is denoted as $\mathbf{d}^M$ as a whole. The scalablity of the interpolation structure lies in the fact that interpolators with different

factors M share the same filter bank $\mathbf{G}(z)$ that is designed offline. Interpolators with different factors M differ only in the multiplication network $\mathbf{d}^M$ that doesn't involve any design strategy. An example of simultaneous implementation of two interpolators with factors M_1 and M_2 is shown in Figure 46.

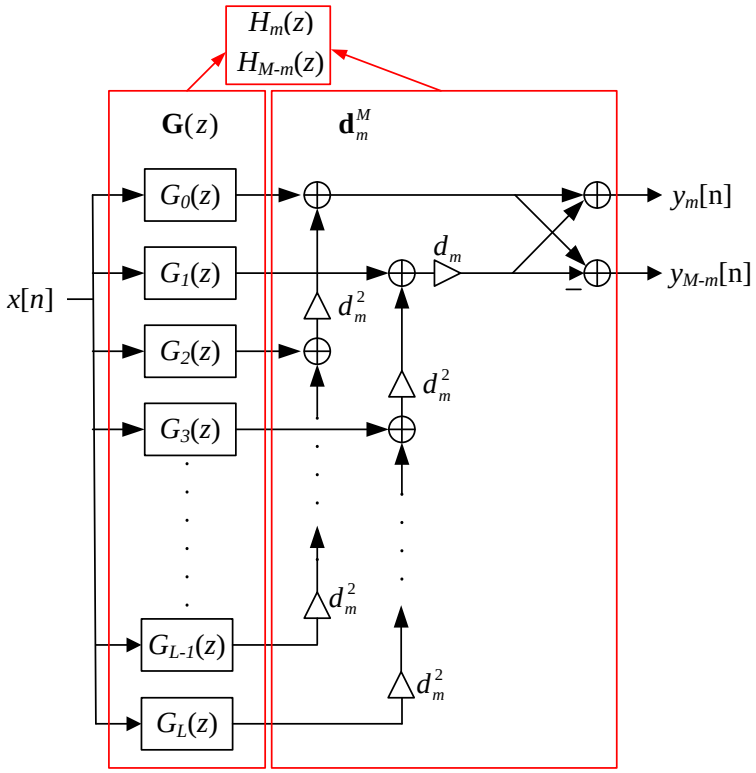

Figure 44: The Farrow structure implementation of $H_m(z)$ and $H_{M-m}(z)$.

2.3.6.1 Design example

The interpolator's magnitude responses using the original Farrow structure with sub-filter order $N_s = 30$ for a variety of interpolation factors M are shown in Figure 47. The corresponding magnitude responses using the Taylor series and optimal polynomial approximated Farrow structures are shown in Figure 48 and 49. The polynomial orders are set as $L = 4$. The Farrow structured FD filter is optimized over $d \in [-0.5, 0.5]$ and $\omega \in [0, 0.95\pi]$. Since the parameter d occupies the entire fractional range of $[-0.5, 0.5]$, the improvement brought by the optimal polynomial

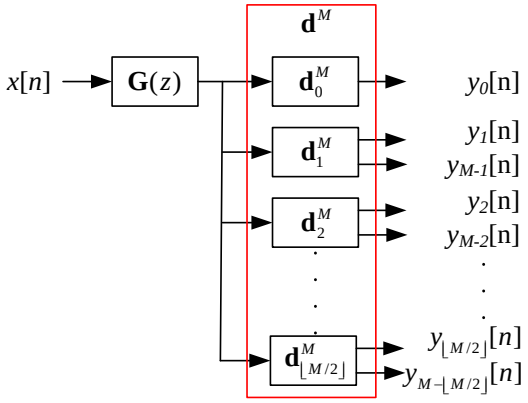

Figure 45: The implementation structure of an M factor interpolator.

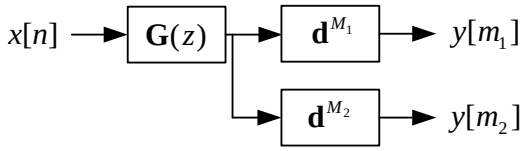

Figure 46: The simultaneous implementation structure of two interpolators.

approximated Farrow structure is relatively small compared with the Taylor series approximation strategy. However, both the Taylor series and optimal polynomial approximation strategies reach a design with much narrower transition band and more consistent attenuation in the stopband.

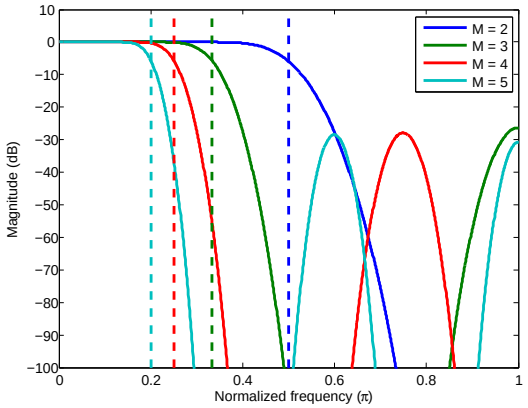

Figure 47: The magnitude response of the interpolator using the original Farrow structure for $N_s = 30$ and $M = [2, 3, 4, 5]$.

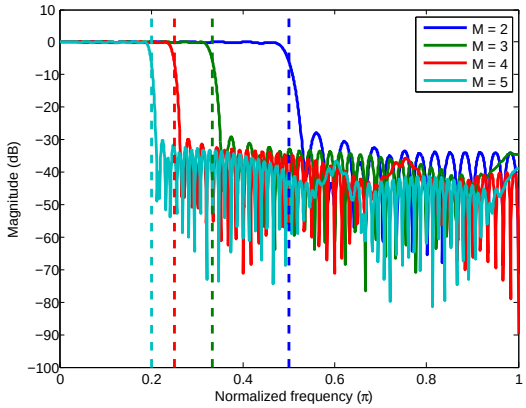

Figure 48: The magnitude response of the interpolator using Taylor series approximated Farrow structure for $N_s = 30$, $L = 4$, $\omega_c = 0.95\pi$ and $M = [2, 3, 4, 5]$.

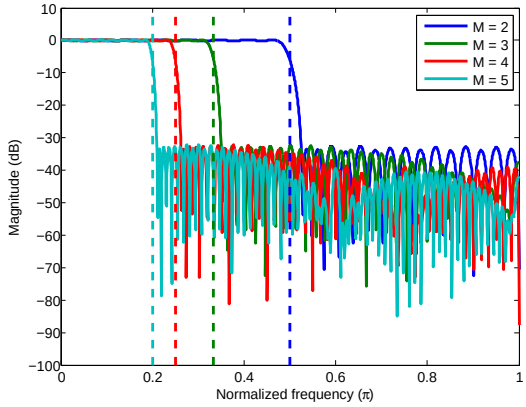

Figure 49: The magnitude response of the interpolator using optimal polynomial approximated Farrow structure for $N_s = 30$, $L = 4$, $\omega_c = 0.95\pi$ and $M = [2, 3, 4, 5]$.

2.3.7 A Filter Structure with Adjustable Cutoff Frequencies Based on the Farrow Structure

Inspired by the scalable multirate filter stucture, [47] proposes the use of the Farrow structure in the realization of a linear phase lowpass filter structure with an adjustable cutoff frequency as shown in Figure 50.

$$H(z) = \sum_{k=0}^{L} G_k(z)(b - b_0)^k. \tag{167}$$

The structure differs from the Farrow structure in that the multipliers are changed from powers of the fractional delay d to the powers of $b - b_0$, where the cutoff frequency b of the resulting lowpass filter is offset by a constant b_0. Although b_0 is optional in the

design, a proper choice of b_0 can reduce the dynamic range of the sub-filter coefficients of $G_k(z)$. Another difference is that all the sub-filters must be the same type linear phase FIR filters in order that the resulting lowpass filter also have linear phase.

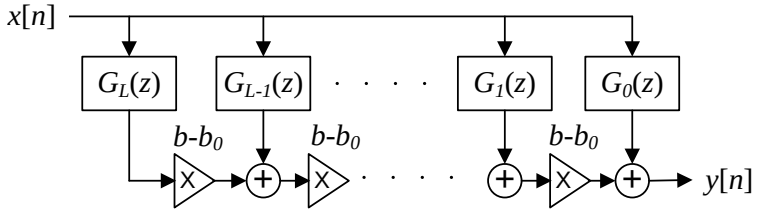

Figure 50: The structure of a lowpass filter structure with adjustable cutoff frequencies.

The design specifications are given by the lower and upper bound of the passband cutoff frequency $[b_l, b_u]$, the ratio of the passband and stopband deviations $r = \frac{\delta_p}{\delta_s}$ and the transition bandwidth Δ. The offset constant is set as $b_0 = \frac{b_l+b_u}{2}$ so that $b - b_0 \in [-\frac{b_u-b_l}{2}, \frac{b_u-b_l}{2}]$. The optimization of the lowpass filter with adjustable cutoff frequencies requires a discretization of the frequency band $[0, \pi]$ and the cutoff frequency range $[b_l, b_u]$. If we denote the resulting filter frequency response at cutoff frequency b_i and frequency ω_j as

$$H^{b_i}(e^{j\omega_j}) = \sum_{k=0}^{L} G_k(e^{j\omega_j})(b_i - b_0)^k. \tag{168}$$

Then the minimax design of the filter is formulated as follows,

$$\begin{aligned} &\min \quad \delta_s \\ &\text{s.t. } |H^{b_i}(e^{j\omega_j}) - 1| \le r\delta_s, \quad \omega_j \in [0, b_i] \\ &\qquad |H^{b_i}(e^{j\omega_j})| \le \delta_s, \quad \omega_j \in [b_i + \Delta, \pi] \end{aligned} \tag{169}$$

The family of magnitude responses for a type I linear phase FIR lowpass filter with adjustable cutoff frequencies is shown in Figure 51. The cutoff frequency range is uniformly discretized into $N_b = 5$ levels. The frequency band is discretized into $N_\omega = 576$ levels. The design algorithm in (169) yields a maximum stopband deviation of $\delta_s = 0.0055$.

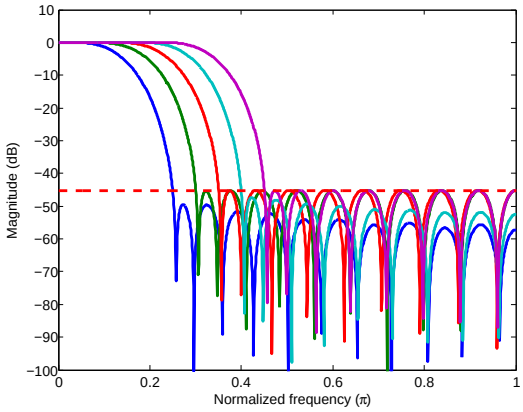

Figure 51: The family of magnitude responses of a lowpass filter with adjustable cutoff frequencies for $N = 24$, $L = 4$, $r = 1$, $b_l = 0.05\pi$, $b_u = 0.45\pi$, $\Delta = 0.1\pi$, $N_b = 5$, $N_\omega = 576$.

The design is not limited to lowpass filters. We can shift the frequency response of a filter by π by multiplying its coefficients with $(-1)^n$. As a result, $H(z)$ is changed to $H(-z)$ and a lowpass filter becomes a high pass filter. Thus a highpass filter with adjustable cutoff frequencies is obtained by multiplying the sub-filter coefficients $g_k[n], n = 0, ..., N_s$ with $(-1)^n$. A bandstop filter can be synthesized by adding a lowpass filter and a highpass filter. We can decompose the sub-filter coefficients $g_k(n)$ according to the parity of n.

$$G_k(z) = G_k^e(z^2) + z^{-1}G_k^o(z^2), \tag{170}$$

where $G_k^o(z^2)$ denotes the frequency response of coefficients of $g_k[n]$ with odd indices of n up-converted by a factor of 2, $G_k^e(z^2)$ denotes the frequency response of the coefficients of $g_k[n]$ with even indices of n up-converted by a factor of 2. It follows

$$G_k(-z) = G_k^e(z^2) - z^{-1}G_k^o(z^2), \tag{171}$$

Denote b^l and b^u as the lower and upper cutoff frequencies of the resulting bandstop

filter, we have

$$H(z) = \sum_{k=0}^{L} [G_k^e(z^2) + z^{-1} G_k^o(z^2)](b^l - b_0)^k + \sum_{k=0}^{L} [G_k^e(z^2) - z^{-1} G_k^o(z^2)](1 - b^u - b_0)^k. \quad (172)$$

At $k = 0$, $[G_0^e(z^2) + z^{-1}G_0^o(z^2)] + [G_0^e(z^2) - z^{-1}G_0^o(z^2)] = 2G_0^e(z^2)$. The resulting bandstop filter with adjustable cutoff frequencies can be implemented as shown in Figure 52. Similarly, we can realize a bandpass filter structure by replacing $2G_0^e(z)$ with $z^{-N/2} - 2G_0^e(z)$. The magnitude response of a bandpass filter with adjustable

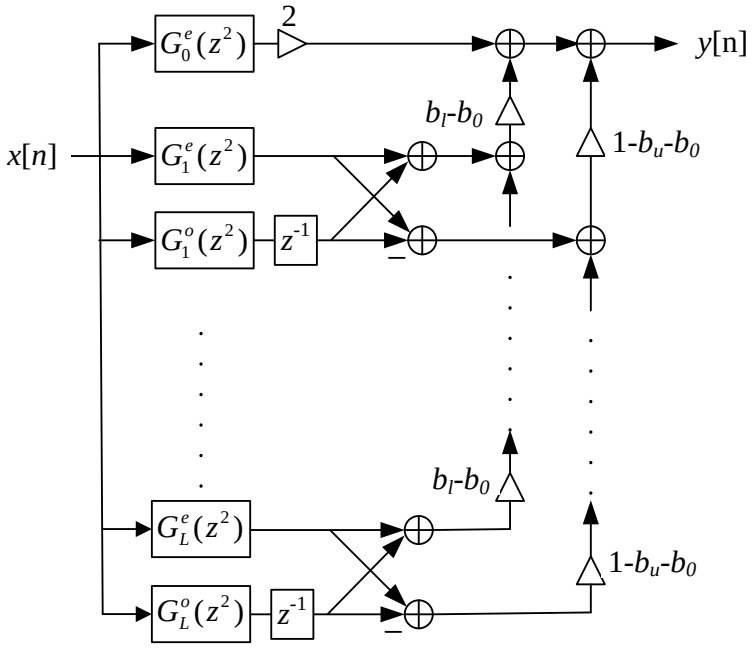

Figure 52: Filter structure for a bandstop filter with adjustable cutoff frequencies.

cutoff frequencies designed with the same specification as in Figure 51 is shown in Figure 53.

2.3.8 A Filter Structure with Adjustable Polynomial Phase Responses Based on the Farrow Structure

The Farrow structured FD filter represents a case where the the slope of the linear phase response is adjustable while the magnitude response is fixed. On the other

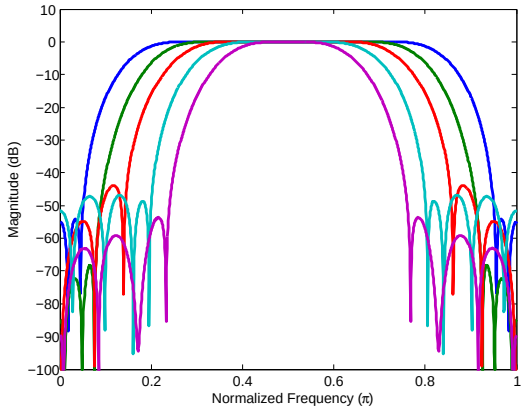

Figure 53: The family of magnitude responses for a bandpass filter with adjustable cutoff frequencies.

hand, a Farrow structured filter structure with adjustable cutoff frequencies represents a case where the linear phase response is fixed while the magnitude response is adjustable. This section generalizes the Farrow structure to realize filters with simultaneously adjustable magnitude and polynomial phases response. Denote the frequency response of an FIR filter of order N as

$$H_{\text{id}}(e^{j\omega}) = \left(\sum_{p=0}^{P} a_p \omega^p\right) e^{-j\left[(N/2)\omega + \sum_{q=1}^{Q} b_q \omega^q\right]}. \tag{173}$$

Note that both the magnitude and phase response of the filter are polynomials in ω. Using the Farrow structure, we can tune both the magnitude and phase responses of the resulting filter by directly adjusting the polynomial coefficients a_p and b_q. By expanding each polynomial phase component into a Taylor series of $jb_q\omega$, we have

$$\begin{aligned} H_{\text{poly}}(e^{j\omega}) &= \left(\sum_{p=0}^{P} a_p \omega^p\right) e^{-j(N/2)\omega} e^{-jb_1\omega} \cdots e^{-jb_Q\omega^Q} \\ &= \left(\sum_{p=0}^{P} a_p \omega^p\right) e^{-j(N/2)\omega} \left(\sum_{k=0}^{L_1} \frac{(-jb_1\omega)^k}{k!}\right) \cdots \left(\sum_{k=0}^{L_Q} \frac{(-jb_Q\omega^Q)^k}{k!}\right) \\ &= e^{-j(N/2)\omega} \sum_{k=0}^{K} c_k \omega^k. \end{aligned} \tag{174}$$

where the complex coefficients c_k are given by the convolution of coefficient coefficients

$$\begin{aligned}\mathbf{c} = &\left[\begin{array}{cccc} a_0 & a_1 & \cdots & a_P \end{array}\right] * \\ &\left[\begin{array}{cccc} 1 & -jb_1 & \frac{(-jb_1)^2}{2!} & \cdots \quad \frac{(-jb_1)^{L_1}}{L_1!} \end{array}\right] * \\ &\left[\begin{array}{ccccccc} 1 & 0 & -jb_2 & 0 & \frac{(-jb_2)^2}{2!} & \cdots & \frac{(-jb_2)^{L_2}}{L_2!} \end{array}\right] * \cdots \\ &\left[\begin{array}{ccccccc} 1 & \underbrace{0 \cdots 0}_{Q-1} & -jb_Q & 0 \cdots 0 & \frac{(-jb_Q)^2}{2!} & \cdots & \frac{(-jb_Q)^{L_Q}}{L_Q!} \end{array}\right]\end{aligned} \tag{175}$$

and

$$K = P + \sum_{q=1}^{Q} qL_q. \tag{176}$$

The approximation error in each polynomial phase component is controlled independently by the Taylor polynomial orders L_q. We can again design $2K$ differentiators of filter order N of distinctive types:

$$\hat{G}_k^R(e^{j\omega}) \to \omega^k e^{-j(N/2)}, \tag{177}$$

$$\hat{G}_k^I(e^{j\omega}) \to j\omega^k e^{-j(N/2)}, \tag{178}$$

which are synthesized by the real and imaginary part of c_k, denoted as c_k^R and c_k^I. The general filter structure is shown in Figure 54.

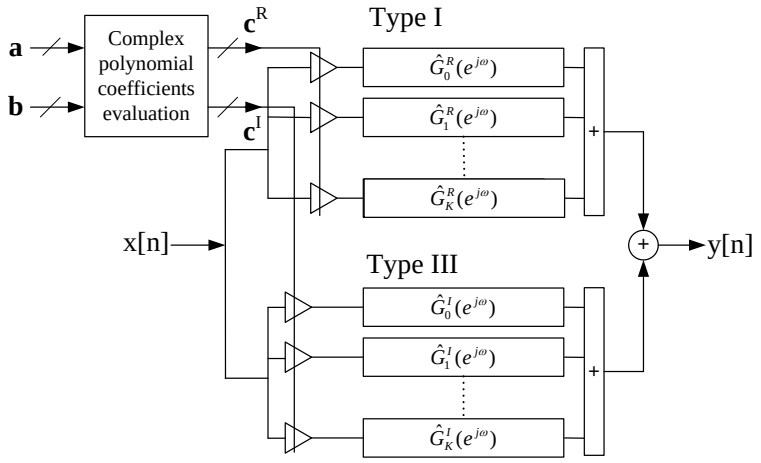

Figure 54: General filter structure with adjustable magnitude and polynomial phase responses.

2.3.8.1 Application

Magnetic resonance imaging (MRI) relies on frequency selective radio-frequency (RF) pulses. The design of such pulses is determined by several criteria, including the excitation profile, the maximum RF field strength $B_{1\,\max}$, the deposited energy and the length of the pulse. For linear-phase pulses, which are essentially sinc-pulses, high frequency selectivity requires a long pulse duration with many side lobes. For a broad excitation bandwidth, pulses of this type require high maximum RF field strength $B_{1\,\max}$, which is ultimately limited by the RF amplifier that drives the magnetic coils. It is shown in [68, 67] that quadratic phase pulses are near-optimal in terms of minimizing the B_1 amplitude for a given bandwidth and flip angle. Previously, the target quadratic-phase FIR filters were designed with a fixed value of for the coefficient of the quadratic phase term using the complex Remez exchange algorithm [38]. With the proposed filter structure in this section, we can make the quadratic coefficient adjustable without redesigning the filters.

Let's look at a concrete example to realize an adjustable quadratic phase filter with a lowpass magnitude response, which means $P = 0$, $a_0 = 1$, $Q = 2$, and $b_1 = 0$.

$$H_{\mathrm{id}}(e^{j\omega}) = \begin{cases} e^{-j\left[(N/2)\omega + b_2\omega^2\right]} & \omega \in [0, \omega_p] \\ 0 & \omega \in [\omega_s, \pi] \end{cases} \tag{179}$$

The passband and stopband cutoff frequencies are given by $\omega_p = 0.095\pi$ and $\omega_s = 0.11\pi$. The range of b_2 is limited to $[-10, 10]$. The filter order is given by $N = 256$. The Taylor polynomial order is set to $L_2 = 3$.

$$\mathbf{c} = \left[1 \quad 0 \quad -jb_2 \quad 0 \quad \tfrac{(-jb_2)^2}{2!} \quad 0 \quad \tfrac{(-jb_2)^3}{3!}\right] \tag{180}$$

The magnitude responses with various values of b_2 are shown in Figure 55. The corresponding phase responses are shown in Figure 56, where the linear phase component $-(N/2)\omega$ has already been removed so that only the quadratic phase component is present.

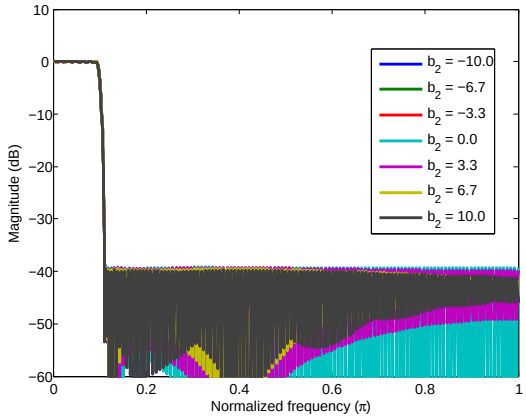

Figure 55: The family of magnitude responses of the adjustable quadratic phase filter.

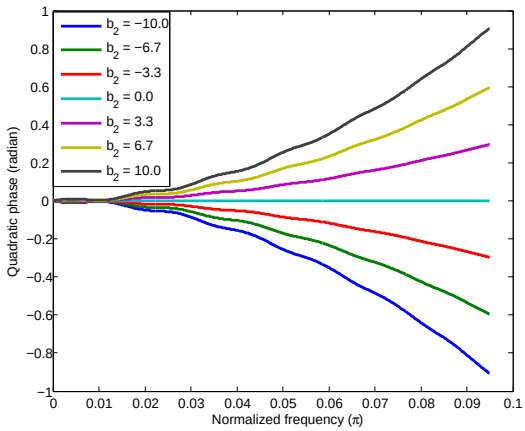

Figure 56: The family of quadratic phase responses of the adjustable quadratic phase filter.

2.4 Multi-coset Sampling of Multi-band Signals

This section deals with another sampling scheme in the recurrent non-uniform sampling category – multi-coset sampling. In TIADC system, there are M channels, each samples at $1/M$ the Nyquist frequency of the input signal. Therefore, the equivalent average sampling frequency still equals to the Nyquist frequency of the input signal. The multi-coset sampling is similar to the TIADC system with the following differences. First, we have L channels where $L < M$, each samples at Ω_s/M so that the equivalent average sampling frequency is $\frac{L}{M}\Omega_s$, which is less than the Nyquist frequency of the signal. Second, we ignore the DC offset and gain mismatches and assume

we have a precise control of the timing mismatches where $\Delta_{t,l} = c_l T_s,\ l = 0, ..., L-1$ and $0 \leq c_l \leq M-1$ is an integer. Third, we are not interleaving the digital output of each channel. Instead, we treat the outputs from different channels as a vector output. Fourth, the input signal is assumed to be a multi-band signal whose band locations are unknown but the band occupancies are sparsely distributed.

Denote the delay integers in each channel as

$$0 \leq c_0 < c_2 < ... < c_{L-1} \leq M-1. \tag{181}$$

Define the l'th sampling sequence as

$$x_l[n] = \begin{cases} x(nT_s) & n = mM + c_l \\ 0 & \text{Otherwise} \end{cases}. \tag{182}$$

An example of the multi-coset sampling strategy is shown in Figure 57

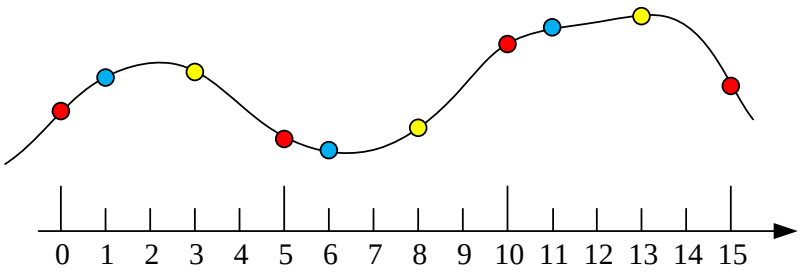

Figure 57: An example of the multi-coset sampling. $M = 5, L = 3, c_0 = 0, c_1 = 1, c_2 = 3$.

The DTFT of $x_l[n]$ is given by

$$X_l(e^{j\Omega T_s}) = \frac{1}{MT_s} \sum_{p=-\infty}^{\infty} e^{j\frac{2\pi}{M}c_l p} X(j(\Omega - p\frac{\Omega_s}{M})). \tag{183}$$

Note that $X(j\Omega)$ is band-limited in $[-\Omega_s/2, \Omega_s/2]$, we have

$$X_l(e^{j\Omega T_s}) = \begin{cases} \frac{1}{MT_s} \sum_{p=-\frac{M-1}{2}}^{\frac{M-1}{2}} e^{j\frac{2\pi}{M}c_l p} X(j(\Omega - p\frac{\Omega_s}{M})),\ \Omega \in [-\frac{\Omega_s}{2M}, \frac{\Omega_s}{2M}) & \text{M odd} \\ \frac{1}{MT_s} \sum_{p=-\frac{M}{2}+1}^{\frac{M}{2}} e^{j\frac{2\pi}{M}c_l p} X(j(\Omega - p\frac{\Omega_s}{M})),\ \Omega \in [0, \frac{\Omega_s}{M}) & \text{M even} \end{cases}. \tag{184}$$

Take odd M case for example, we can write the above equation in matrix from:

$$\mathbf{x}(e^{j\Omega T_s}) = \mathbf{A}\mathbf{x}(j\Omega), \quad \text{for } \Omega \in [-\frac{\Omega_s}{2M}, \frac{\Omega_s}{2M}] \tag{185}$$

where

$$\mathbf{x}(e^{j\Omega T_s}) = \left[\begin{array}{cccc} X_0(e^{j\Omega T_s}) & ,..., & X_{L-1}(e^{j\Omega T_s}) \end{array} \right]^T, \tag{186}$$

$$\mathbf{x}(j\Omega) = \left[\begin{array}{cccc} X(j(\Omega - \frac{M-1}{2M}\Omega_s)) & ,..., & X(j(\Omega + \frac{M-1}{2M}\Omega_s)) \end{array} \right]^T, \tag{187}$$

$$\mathbf{A}_{l,m} = \frac{1}{MT_s} e^{j\frac{2\pi}{M}c_l(m-\frac{M-1}{2})}. \tag{188}$$

This is an under-determined system of equations as long as $L < M$. However, $\mathbf{X}(j\Omega)$ is still recoverable as long as its band occupancies are sparse and the sub-DFT matrix $\mathbf{A}$ is well designed. [64] proved the existence of a universal sampling pattern, which guaranteed the reconstruction from samples as long as the signals has a spectral occupancy rate lower than a given bound. The frequency Ω in (185) is a continuous quantity. We have to quantize Ω in order to solve (185). Suppose we uniformly quantize Ω into N grid in $[-\frac{\Omega_s}{2M}, \frac{\Omega_s}{2M}]$, we can obtain the following matrix representation of $X(j\Omega)$ as shown in Figure 58. Denote

$$\mathbf{\Omega} = [-\frac{\Omega_s}{2M}, \cdots, -\frac{\Omega_s}{2M} + \frac{n\Omega_s}{MN}, \cdots, \frac{\Omega_s}{2M} - \frac{\Omega_s}{MN}]^T, \tag{189}$$

equation (185) can be re-written as

$$\mathbf{X}(e^{j\mathbf{\Omega}T_s}) = \mathbf{A}\mathbf{X}(j\mathbf{\Omega}), \tag{190}$$

where $\mathbf{X}(e^{j\mathbf{\Omega}T_s})$ is a matrix of dimension L by N, whose (l,n)-th entry is given by $X_l(e^{j\Omega_n T_s})$. Similarly, $\mathbf{X}(j\mathbf{\Omega})$ is a matrix of dimension M by N, whose (m,n)-th entry is given by $X(j(\Omega_n - (\frac{M-1}{2} - m)\frac{\Omega_s}{M})$. Equation (190) is a multiple measurement vector (MMV) [17, 15] problem [17, 15] in compressive sensing. The sparse multi-band signal assumption implies that the columns in the solution matrix $\mathbf{X}(j\mathbf{\Omega})$ shares the same support whose cardinality is much less than M. Although equation (190) represents an under-determined system of equations. We can still recover $\mathbf{X}(j\mathbf{\Omega})$ from $\mathbf{X}(e^{j\mathbf{\Omega}T_s})$ should $\mathbf{X}(j\mathbf{\Omega})$ satisfy the sparse multi-band assumption by solving the MMV problem.

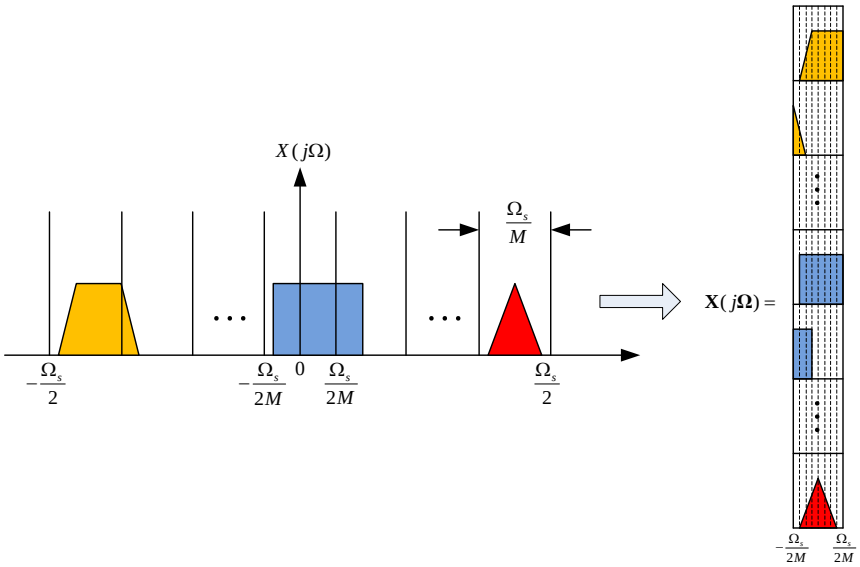

Figure 58: The matrix representation of $X(j\Omega)$

Chapter III

NON-RECURRENT NON-UNIFORM SAMPLING

The TIADC and multi-coset sampling systems discussed in Chapter 2 are recurrent in a sense that the non-uniformities repeat themselves periodically. In the TIADC system, the non-uniformities are unknown and need to be calibrated. The equivalent sampling rate of the TIADC system is still above the Nyquist rate of the input signal. The multi-coset sampling system, on the other hand, assumes that the channel timing mismatches are precisely controllable. The input signal is assumed to be a multi-band signal whose frequency band occupancy is sparse. Such a signal can be reconstructed even when sampled below its Nyquist rate according to the multi-coset sampling scheme.

This chapter deals with the other class of non-uniform sampling – non-recurrent and non-uniform sampling. The sampling process doesn't show any periodicity and can only be described statistically. The theoretical framework is based on the random sampling theory established by Beutler and Leneman in the 60s. However, a different perspective is provided here to interpret the aliases as a noise source. The effect of time quantization is analyzed from a theoretical perspective which leads to several practical implementations of the non-recurrent and non-uniform sampling system. Unlike uniform sampling, it is shown in this chapter that the Nyquist rate is no longer the barrier under a random sampling scheme. It is possible to unambiguously reconstruct a class of spectrally sparse signals that is sampled randomly below the Nyquist rate.

3.1 Spectral Analysis of Random Impulse Processes

This section deals with the situation where we have precise knowledge and control over the non-uniformities in the sampling process. A theoretical analysis is given to associate the probability distribution of the sampling time points with the aliasing effect of the sampling process. A variety of random sampling models are investigated to achieve aliasing suppression and sub-Nyquist rate sampling.

Definition 1. *A random impulse process $s(t)$ is defined as*

$$s(t) = \sum_{n=-\infty}^{\infty} \alpha_n \delta(t - t_n), \tag{191}$$

where $\{t_n\}$ is a stationary point process (SPP), $\{\alpha_n\}$ is a discrete time stationary random process independent of $\{t_n\}$.

Suppose that a random process $x(t)$ is sampled by a random impulse process $s(t)$ that is independent of $x(t)$,

$$y(t) = x(t)s(t). \tag{192}$$

In the frequency domain, we have

$$\Phi_y(f) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi_x(f-u)\Phi_s(u)du = \Phi_x(f) * \Phi_s(f), \tag{193}$$

where $\Phi_y(f)$, $\Phi_x(f)$, $\Phi_s(f)$ stand for the power spectral densities (PSD) of $y(t)$, $x(t)$ and $s(t)$, respectively. If $\Phi_s(f)$ contains a single impulse $\delta(f)$, then no aliasing exists in the output spectrum. We can analyze the aliasing suppression capability of various random impulse processes by studying their PSDs $\Phi_s(f)$.

Beutler and Leneman established a systematic theory for random sampling [3, 4, 5, 6] based on the analytic expression of $\Phi_s(f)$. This section first introduces some basic definitions and results from Beutler and Leneman's work and then derives a series of results for various random sampling models.

First, the stationary point process is defined as follows.

Definition 2. *Let* $\{\tau_n\}, n = 0, \pm 1, \pm 2, \cdots$ *be a discrete time random process such that* $\tau_n \geq 0$ *for all* n *and all* τ_n *are finite valued. If we have*

$$t_n = \begin{cases} \tau_0 + \sum_{k=1}^{n} \tau_k & n \geq 1 \\ \tau_0 & n = 0 \\ \tau_0 - \sum_{k=n}^{-1} \tau_k & n \leq -1 \end{cases},$$

then $\{t_n\}$ *is called a random point process.*

A random point process is stationary if the number of points in a set of fixed time intervals is invariant under any time shift which preserves the length and spacing of those intervals. Written mathematically, denote $A_n(t, x)$ as the event "there are exactly n points in $(t, t + x]$", we have the following definition for stationary point process:

Definition 3. $\{t_n\}$ *is a stationary point process (SPP) if for any intervals* $x_1, x_2, \cdots x_n$, *time points* $t_1, t_2, \cdots t_n$, *numbers of points* $k_1, k_2, \cdots k_n$ *and time shift* h,

$$Pr\left[\bigcap_{j=1}^{n} A_{k_j}(t_j, x_j)\right] = Pr\left[\bigcap_{j=1}^{n} A_{k_j}(t_j + h, x_j)\right],$$

where $Pr[\cdot]$ *represents the probability of the corresponding event.*

For example, if the τ_k's are mutually independent and exponentially distributed with the exponential probability density function (PDF)

$$f_{\tau_k}(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{Otherwise} \end{cases}, \tag{194}$$

then $\{t_n\}$ is a SPP, which is referred as a Poisson SPP. In another example, if $\tau_k = T$, which corresponds to the uniform sampling case, $\{t_n\}$ is also an SPP, which is referred as a periodic SPP. Beutler and Leneman established three theorems following their definition of stationary point process.

Theorem 1. *If $\{t_n\}$ is an SPP, $\{x_n\}$ is a discrete time stationary random process that takes the value of $\{0, 1\}$ and is independent of $\{t_n\}$, the new point process is formulated as follows: a new point process has a point at $\{t_n\}$ if $x_n = 1$ but no point at $\{t_n\}$ if $x_n = 0$, then the new point process $\{t'_n\}$ is also an SPP, referred as random skip SPP.*

Corollary 1. *Suppose $Pr[x_n = 0] = q$, denote the characteristic function (CF) of the interval τ_k as $\psi_{\tau_k}(f)$ with the following definition:*

$$\psi_{\tau_k}(f) = E[e^{jf\tau_k}] = \int_{-\infty}^{\infty} e^{jfx} f_{\tau_k}(x)dx, \tag{195}$$

then the characteristic function of the new interval τ'_k is given by

$$\psi_{\tau'_k}(f) = \frac{(1-q)\psi_{\tau_k}(f)}{1 - q\psi_{\tau_k}(f)}. \tag{196}$$

Proof. We can relate the characteristic function to the Fourier transform of the PDF of τ_k:

$$\psi_{\tau_k}(-2\pi f) = \mathcal{F}\{f_{\tau_k}(x)\}, \tag{197}$$

where $\mathcal{F}\{\cdot\}$ stands for the continuous time Fourier transform (CTFT).

$$\begin{aligned}
\psi_{\tau'_k}(-2\pi f) =& \mathcal{F}\{(1-q)f_{\tau_k}(x) + q(1-q)f_{\tau_k}(x) * f_{\tau_k}(x) && (198)\\
&+ q^2(1-q)f_{\tau_k}(x) * f_{\tau_k}(x) * f_{\tau_k}(x) + \cdots\} && (199)\\
=& (1-q)\sum_{n=1}^{\infty} q^{n-1}\mathcal{F}\{f_{\tau_k}(x)\}^n && (200)\\
=& \frac{(1-q)\psi_{\tau_k}(-2\pi f)}{1 - \psi_{\tau_k}(-2\pi f)} && (201)
\end{aligned}$$

□

Theorem 2. *If $\{t_n\}$ is a periodic SPP with period T, $\{u_n\}$ is a discrete time stationary random process with $-T/2 \le u_n < T/2$, then the new point process $t'_n = t_n + u_n$ is also an SPP, which is referred as jittered SPP.*

The most important results of the random sampling theory is described in the following theorem, which relates the distribution of sampling intervals τ_k with the PSD of the sampling process $s(t)$.

Theorem 3. *Denote $\beta = \frac{1}{E[\tau_k]}$ as the average points per unit time interval, the autocorrelation function of α_n as $R_\alpha[n]$, the sum of n consecutive intervals as*

$$\sigma_n = \sum_{k=j}^{j+n-1} \tau_k \qquad \forall j, \tag{202}$$

the characteristic function of σ_n as $\psi_{\sigma_n}(f)$, then the autocorrelation function of $s(t)$ is given by

$$R_s(\tau) = \beta R_\alpha[0]\delta(\tau) + \beta \sum_{n=1}^{\infty} R_\alpha[n] f_{\sigma_n}(|\tau|), \tag{203}$$

Accordingly, the PSD of $s(t)$ is given by

$$\Phi_s(f) = \beta R_\alpha[0] + 2\beta \sum_{n=1}^{\infty} \Re\left[R_\alpha[n]\psi_{\sigma_n}(2\pi f)\right]. \tag{204}$$

Proof.

$$\begin{aligned}
\Phi_s(f) =& \beta R_\alpha[0] + \beta \sum_{n=1}^{\infty} R_\alpha[n] \int_{-\infty}^{\infty} f_{\sigma_n}(|\tau|) e^{-j2\pi f\tau} d\tau & (205)\\
=& \beta R_\alpha[0] + \beta \sum_{n=1}^{\infty} R_\alpha[n] \left[\int_0^{\infty} f_{\sigma_n}(\tau) e^{-j2\pi f\tau} d\tau + \int_0^{\infty} f_{\sigma_n}(\tau) e^{j2\pi f\tau} d\tau\right] & (206)\\
=& \beta R_\alpha[0] + \beta \sum_{n=1}^{\infty} R_\alpha[n] \left[\psi_{\sigma_n}(2\pi f) + \psi_{\sigma_n}(-2\pi f)\right] & (207)
\end{aligned}$$

Since σ_n is real, we have

$$\psi_{\sigma_n}(2\pi f) = \psi^*_{\sigma_n}(-2\pi f) \tag{208}$$

$$\Phi_s(f) = \beta R_\alpha[0] + 2\beta \sum_{n=1}^{\infty} \Re\left[R_\alpha[n]\psi_{\sigma_n}(2\pi f)\right] \tag{209}$$

□

If we further assume that α_n is exponentially correlated such that

$$R_\alpha[n] = \rho^{|n|}, \tag{210}$$

where $-1 \leq \rho \leq 1$, (204) can be simplified to

$$\Phi_s(f) = \beta \Re \left[\frac{1 + \rho \psi_{\tau_k}(2\pi f)}{1 - \rho \psi_{\tau_k}(2\pi f)} \right]. \tag{211}$$

Such a random amplitude process α_n can be obtained by filtering a zero mean, unit variance white noise v_n through a first order autoregressive model:

$$\alpha_n = \rho \alpha_{n-1} + \sqrt{1 - \rho^2} v_n. \tag{212}$$

Alternatively, α_n can be generated according to a binary, zero mean Markov process according to

$$\Pr[\alpha_{n+1} \neq \alpha_n] = p, \tag{213}$$

$$\Pr[\alpha_{n+1} = \alpha_n] = 1 - p, \tag{214}$$

whose autocorrelation function is

$$R_\alpha[n] = (1 - 2p)^{|n|}. \tag{215}$$

When the amplitude process follows a Bernoulli distribution:

$$\Pr[\alpha_n] = \begin{cases} p & \alpha_n = 1 \\ 1 - p & \alpha_n = -1 \end{cases}, \tag{216}$$

its autocorrelation function becomes

$$R_\alpha[n] = \begin{cases} 1 & n = 0 \\ (2p-1)^2 & n \neq 0 \end{cases}. \tag{217}$$

Equation (204) becomes

$$\Phi_s(f) = \beta + 2\beta(2p-1)^2 \sum_{n=1}^{\infty} \Re \left[\psi_{\sigma_n}(2\pi f) \right]. \tag{218}$$

When $\rho = 1$, we have $\alpha_n \equiv 1$. The random sampling process degenerates into a non-uniformly spaced impulse train. From (204), we can conclude that there is a

broadband additive noise term $\beta R_\alpha[0]$ in $\Phi_s(f)$, which can be interpreted as the aliasing noise. From (203), a sufficient condition for $\Phi_s(f)$ to be aliasing frequency free when $\rho = 1$ is that

$$f_s(t) = \sum_{n=1}^{\infty} f_{\sigma_n}(t) = \text{const.}, \tag{219}$$

where $f_s(t)$ is a sampling point density function. Assuming τ_k is continuous and independently and identically distributed (IID) with mean $1/\beta$, then

$$f_s(t) = f_{\tau_k}(t) + f_{\tau_k}(t) * f_{\tau_k}(t) + f_{\tau_k}(t) * f_{\tau_k}(t) * f_{\tau_k}(t) + \cdots \tag{220}$$

According to the large number theorem,

$$f_s(t \to \infty) = \beta. \tag{221}$$

Therefore, any continuous and integrable IID interval distribution makes the sampling process aliasing free after a certain time. The convergence speed of $f_s(t)$ toward β is distribution dependent.

Figures 59 and 60 show the convergence speed of $f_s(t)$ for uniformly and exponentially distributed intervals. In the uniformly distributed interval case, $f_s(t)$ converges to β after approximately 0.5 seconds while $f_s(t)$ in the exponentially distributed interval case is consistently β. σ_n converges to a Gaussian distribution as $n \to \infty$ regardless the distribution of τ_k.

3.1.1 Additive Random Sampling (ARS)

Under ARS, the sampling time t_k is given by

$$t_k = t_{k-1} + \tau_k, \tag{222}$$

where τ_k's are independently and identically distributed (IID).

3.1.1.1 Exponentially distributed intervals

Consider the exponentially distributed interval case where $\beta = \lambda$:

$$\psi_{\tau_k}(2\pi f) = \frac{\lambda}{\lambda - j2\pi f}. \tag{223}$$

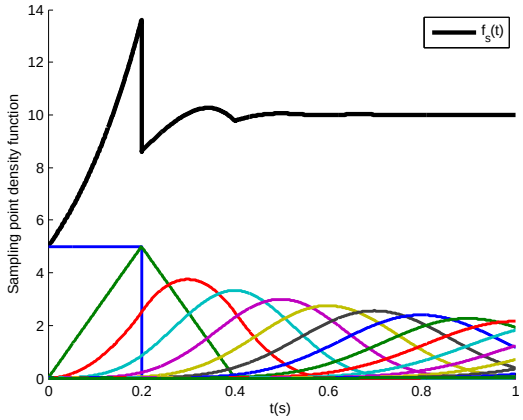

Figure 59: $f_s(t)$ for uniformly distributed intervals, $\tau_k \sim U[0, 0.2]s$.

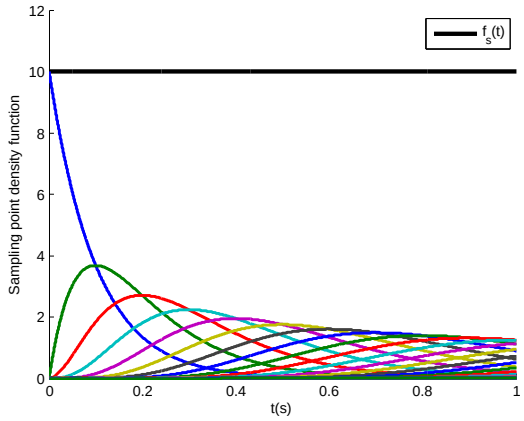

Figure 60: $f_s(t)$ for exponentially distributed intervals, $\lambda = 10$.

When $|\rho| < 1$,

$$\Phi_s(f) = \lambda \frac{\lambda^2(1-\rho^2) + (2\pi f)^2}{\lambda^2(1-\rho)^2 + (2\pi f)^2}. \tag{224}$$

$$\Phi_s(0) = \lambda \frac{1+\rho}{1-\rho}. \tag{225}$$

When $\rho = 1$,

$$\Phi_s(f) = \lambda^2 \delta(f) + \lambda. \tag{226}$$

Proof. Since each τ_k is exponentially distributed with parameter λ, σ_n follows an Erlang distribution with parameters (n, λ):

$$f_{\sigma_n}(x) = \begin{cases} \frac{\lambda^n x^{n-1} e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & \text{Otherwise} \end{cases}. \tag{227}$$

(203) becomes

$$R_s(\tau) = \lambda\delta(\tau) + \lambda \sum_{n=1}^{\infty} \frac{\lambda^n \tau^{n-1} e^{-\lambda\tau}}{(n-1)!} \tag{228}$$

$$= \lambda\delta(\tau) + \lambda^2 e^{-\lambda\tau} \sum_{n=0}^{\infty} \frac{\lambda\tau^n}{n!} \tag{229}$$

$$= \lambda\delta(\tau) + \lambda^2 \tag{230}$$

Therefore,

$$\Phi_s(f) = \lambda^2\delta(f) + \lambda \tag{231}$$

□

Therefore, ARS with exponentially distributed intervals is aliasing free since there is only a single impulse at $f = 0$. The additive term λ can be viewed as additive white noise. Unlike uniform sampling, which results in the presence of regularly spaced aliases, exponentially distributed ARS shapes those aliases into white noise with a flat spectrum over $-\infty < f < \infty$. Suppose $x(t)$ is an analytic signal with $\Phi_x(f) = A^2\delta(f - f_0)$, then $\Phi_y(f) = A^2\lambda^2\delta(f - f_0) + A^2\lambda$. We can evaluate the normalized power spectrum $P_y(f)$ according to

$$P_y(f) = \frac{1}{N^2} \left| \sum_{n=0}^{N-1} y(t_n) e^{-j2\pi f t_n} \right|^2, \tag{232}$$

where N is the number of non-uniform samples. For the above mentioned example,

$$P_y(f) \to A^2\delta(f - f_0) + A^2/N, \tag{233}$$

which indicate that we can reduce the aliasing noise floor A^2/N by taking more samples. When applying random skip to exponentially distributed ARS, we have

$$\psi_{\tau_k'}(2\pi f) = \frac{(1-q)\lambda}{(1-q)\lambda - j2\pi f}. \tag{234}$$

We only need to replace the λ with $(1 - q)\lambda$.

Figure 61 shows the shape of $\Phi_s(f)$ as ρ changes from 0.1 to 0.9. $\Phi_s(f)$ approximates an impulse function at $f = 0$ as ρ approximates 1. Figure 62 shows the power spectra of a sampled analytic signal $x(t) = e^{j2\pi f_0 t}$, where $f_0 = 5$ Hz. The number of samples is $N = 1024$. The average sampling frequency is $\lambda = 3$ Hz. The uniform sampling case has a sampling frequency $F_s = \lambda$. Accordingly, the power spectrum of the uniformly sampled signals is only shown in the range of $[-F_s/2, F_s/2]$. The random amplitude process $\{\alpha_n\}$ has a correlation coefficient of $\rho = 0.9$. In the uniform sampling case, since $F_s/2 < f_0$, an aliasing frequency at $f_0 - 2F_s = -1$ Hz is shown in the baseband. While in the exponentially distributed ARS case, no aliasing is observed. The aliasing noise floor converges to $1/N$ as f gets further away from f_0. Figure 63 shows the signal power spectra similar to the case in Figure 62. The only difference is that $\rho = 1$ in this case. The aliasing noise floor is flat.

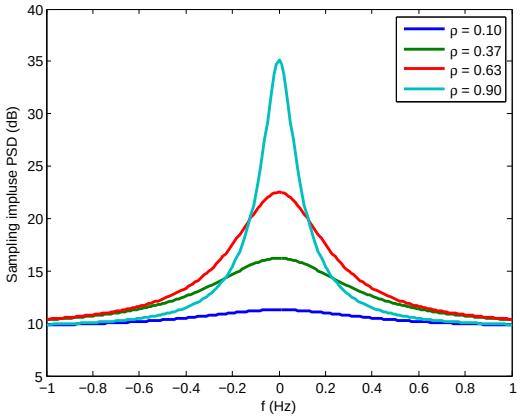

Figure 61: The shape of $\Phi_s(f)$ as a function of ρ for exponentially distributed ARS. $\lambda = 3$ Hz.

3.1.1.2 Uniformly distributed intervals

Consider the case of uniformly distributed intervals. Suppose τ_k is uniformly distributed in $[a, b]$, where $a \geq 0$, $\beta = \frac{1}{E[\tau_k]} = \frac{2}{a+b}$.

$$\psi_{\tau_k}(2\pi f) = e^{j\pi f(a+b)} \frac{\sin\left[2\pi f\left(\frac{b-a}{2}\right)\right]}{2\pi f\left(\frac{b-a}{2}\right)}. \tag{235}$$

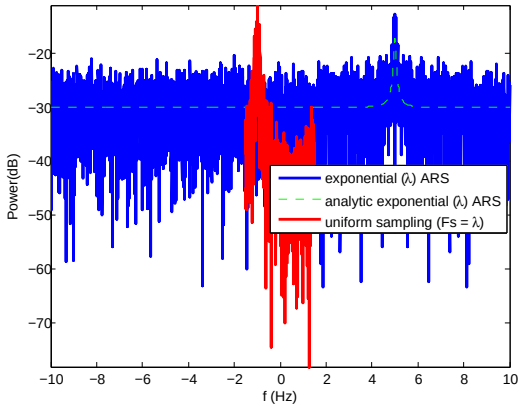

Figure 62: The power spectra of a sampled analytic signal with a frequency at 5 Hz for exponentially distributed ARS. $\lambda = 3$ Hz, $\rho = 0.9$, N $= 1024$.

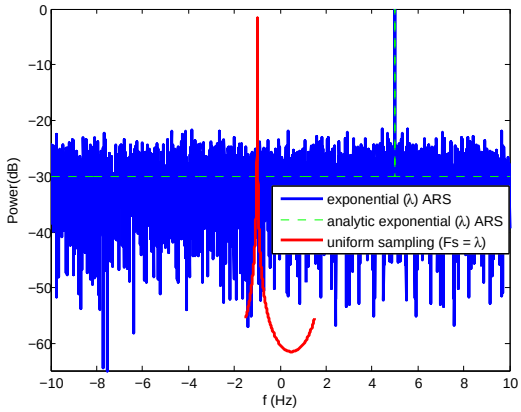

Figure 63: The power spectra of a sampled analytic signal with a frequency at 5 Hz for exponentially distributed ARS. $\lambda = 3$ Hz, $\rho = 1$, N $= 1024$.

When $|\rho| < 1$,

$$\Phi_s(f) = \frac{2}{a+b} P(\rho \frac{\sin(\pi f(b-a))}{\pi f(b-a)}, \pi f(b+a)) \tag{236}$$

$$= \frac{2}{a+b} \frac{1 - \left(\rho \frac{\sin(\pi f(b-a))}{\pi f(b-a)}\right)^2}{1 - 2\rho \frac{\sin(\pi f(b-a))\cos(\pi f(b+a))}{\pi f(b-a)} + \left(\rho \frac{\sin(\pi f(b-a))}{\pi f(b-a)}\right)^2}, \tag{237}$$

where $P(r,\theta)$ is a Poisson kernel with the following definition:

$$P(r,\theta) = \sum_{n=-\infty}^{\infty} r^{|n|} e^{jn\theta} \tag{238}$$

$$= \Re\left\{\frac{1+re^{j\theta}}{1-re^{j\theta}}\right\} \tag{239}$$

$$= \frac{1-r^2}{1-2r\cos\theta+r^2}, \tag{240}$$

where, $0 \leq r < 1$. Poisson kernel is periodic in θ:

$$P(r,\theta) = P(r,\theta+2\pi n), \quad \forall n \in \mathbb{Z}, \tag{241}$$

and

$$\lim_{r\to 1} P(r,\theta) = \sum_{n=-\infty}^{\infty} e^{jn\theta} = 2\pi \sum_{n=-\infty}^{\infty} \delta(\theta - 2\pi n). \tag{242}$$

When $\rho = 1$,

$$\Phi_s(f) = \begin{cases} \frac{2}{a+b}\frac{1-\left(\frac{\sin(\pi f(b-a))}{\pi f(b-a)}\right)^2}{1-2\frac{\sin(\pi f(b-a))\cos(\pi f(b+a))}{\pi f(b-a)}+\left(\frac{\sin(\pi f(b-a))}{\pi f(b-a)}\right)^2} & f \neq 0 \\ \left(\frac{2}{a+b}\right)^2 \delta(f) & f = 0 \end{cases} \tag{243}$$

Proof. We only need to concentrate on the case where $f = 0$ which makes $r = 1$,

$$\Phi_s(f) = \frac{2}{a+b} 2\pi\delta(\pi f(a+b)) = \left(\frac{2}{a+b}\right)^2 \delta(f). \tag{244}$$

□

Like the exponentially distributed interval case, ARS with uniformly distributed intervals is also aliasing free. However, the aliasing noise floor is not flat. Figure 64 shows the shape of $\Phi_s(f)$ as ρ changes from 0.1 to 0.9. $\Phi_s(f)$ approximates an impulse function at $f = 0$ as ρ approximates 1. Figure 65 shows the power spectra of a sampled analytic signal $x(t) = e^{j2\pi f_0 t}$, where $f_0 = 5$ Hz. The number of samples is $N = 1024$. The average sampling rate is $\frac{2}{a+b} = 3$ Hz. The uniform sampling case also has a sampling frequency $F_s = 3$ Hz. Accordingly, the power spectrum of the uniformly sampled signal is only shown in the range of $[-F_s/2, F_s/2]$. The random

amplitude process has a correlation coefficient of $\rho = 0.9$. In the uniform sampling case, an aliasing frequency $f_0 - 2F_s = -1$ Hz is shown in the Nyquist range. While in the uniformly distributed ARS case, no aliasing frequencies are observed excerpt for a non-flat aliasing noise floor. Figure 66 shows the signal power spectra similar to the case in Figure 65. The only difference is that $\rho = 1$ in this case. The noise floor is also non-flat.

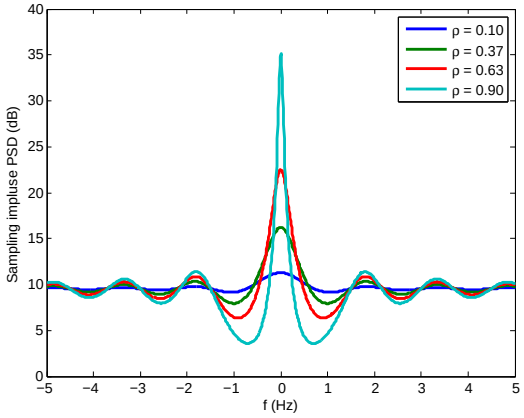

Figure 64: The shape of $\Phi_s(f)$ as a function of ρ for uniformly distributed ARS. a = 0 s, $b = 2/3$ s.

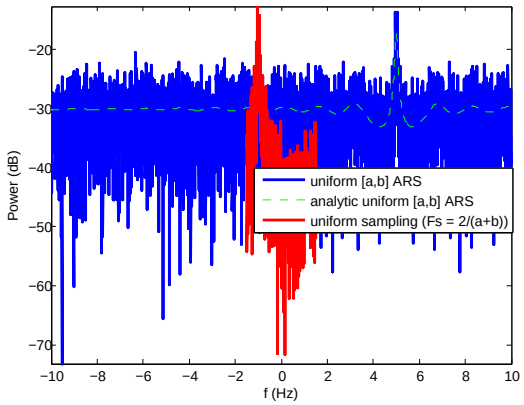

Figure 65: The power spectra of a sampled analytic signal with a frequency at 5 Hz for uniformly distributed ARS. $a = 0$ s, $b = 2/3$ s, $\rho = 0.9$, N = 1024.

3.1.2 Jittered Random Sampling (JRS)

Under JRS, IID jitters $u_k \in [-T/2, T/2]$ are applied to a uniform sampling grid $t_k = kT$ so that new sampling time is $t'_k = kT + u_k$. The sampling interval is given

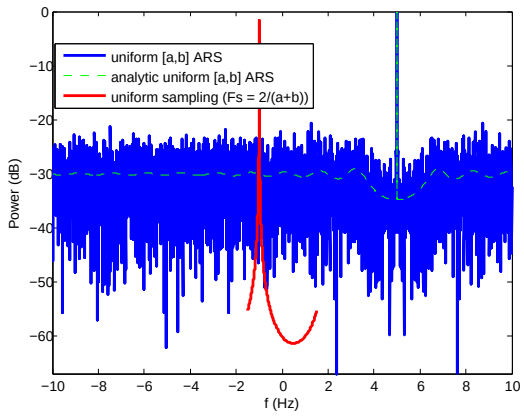

Figure 66: The power spectra of a sampled analytic signal with a frequency at 5 Hz for uniformly distributed ARS. $a = 0$ s, $b = 2/3$ s, $\rho = 1$, N = 1024.

by

$$\tau_k = t'_k - t'_{k-1} \tag{245}$$

$$= t_k - t_{k-1} + u_k - u_{k-1} \tag{246}$$

$$= T + u_k - u_{k-1}. \tag{247}$$

Therefore,

$$\sigma_n = \sum_{k=j}^{n+j-1} \tau_k = nT + u_{n+j-1} - u_{j-1}, \quad \forall j. \tag{248}$$

$$\psi_{\sigma_n}(2\pi f) = E\left[e^{j2\pi f \sigma_n}\right] = e^{j2\pi nfT} |\psi_{u_k}(2\pi f)|^2, \tag{249}$$

where $\psi_{u_k}(f)$ is the characteristic function of u_k. When combined with random skip,

$$\psi_{\sigma_n}(2\pi f) = \frac{(1-q)e^{j2\pi nfT}}{1 - qe^{j2\pi nfT}} |\psi_{u_k}(2\pi f)|^2. \tag{250}$$

When $|\rho| < 1$,

$$\Phi_s(f) = \frac{1-q}{T}\left\{1 + 2\sum_{n=1}^{\infty} \Re\left[\rho^n \frac{(1-q)e^{j2\pi nfT}}{1 - qe^{j2\pi nfT}} |\psi_{u_k}(2\pi f)|^2\right]\right\} \tag{251}$$

$$= \frac{1-q}{T}\left\{1 + \frac{\rho(1-q)}{\mu} |\psi_{u_k}(2\pi f)|^2 \Re\left[\frac{1 + \mu e^{j2\pi fT}}{1 - \mu e^{j2\pi fT}} - 1\right]\right\} \tag{252}$$

$$= \frac{1-q}{T}\left\{1 - \frac{\rho(1-q)}{\mu} |\psi_{u_k}(2\pi f)|^2 + \frac{\rho(1-q)}{\mu} |\psi_{u_k}(2\pi f)|^2 P(\mu, 2\pi fT)\right\}, \tag{253}$$

where $\mu = \rho - \rho q + q \in (-1, 1)$. When $\rho = 1$, we have $\mu = 1$,

$$\Phi_s(f) = \frac{1-q}{T}\left\{1 - (1-q)|\psi_{u_k}(2\pi f)|^2 + \frac{1-q}{T}\sum_{n=-\infty}^{\infty}|\psi_{u_k}(\frac{2\pi n}{T})|^2\delta(f - \frac{n}{T})\right\}. \tag{254}$$

Generally speaking, JRS is not aliasing frequency free. The n-th aliasing term is scaled by a factor of $|\psi_{u_k}(\frac{2\pi n}{T})|^2$. There is also a aliasing noise term $1 - \frac{\rho(1-q)}{\mu}|\psi_{u_k}(2\pi f)|^2$ in the power spectrum.

3.1.2.1 Uniformly distributed jitters

Assuming u_k is uniformly distributed in $[-\frac{a}{2}, \frac{a}{2}]$,

$$\psi_{u_k}(2\pi f) = \frac{\sin(\pi f a)}{\pi f a}. \tag{255}$$

Each aliasing term is scaled by $\left[\frac{\sin(\pi n a/T)}{\pi n a/T}\right]^2$. When $a = T$, the aliasing scaling term becomes

$$\psi_{u_k}(2\pi\frac{n}{T}) = \frac{\sin(\pi n)}{\pi n} = \begin{cases} 1 & n = 0 \\ 0 & n \neq 0 \end{cases}. \tag{256}$$

$$\Phi_s(f) = \frac{1-q}{T}\left\{1 - (1-q)\left(\frac{\sin(\pi f T)}{\pi f T}\right)^2 + \frac{1-q}{T}\delta(f)\right\}. \tag{257}$$

Therefore, uniform JRS is aliasing frequency free if the jitters u_k are uniformly distributed in $[-T/2, T/2]$.

Figure 67 shows the shape of $\Phi_s(f)$ as ρ changes from 0.1 to 0.9. $a = T/2$, $\Phi_s(f)$ shows multiple impulses at $f = \frac{n}{T}$ which are scaled by $\left(\frac{\sin(\pi f a)}{\pi f a}\right)^2$ (shown in the dashed lines) as ρ approximates 1. Figure 68 shows the case where $a = T$. The aliasing impulses happen to be scaled by the nulls of $\left(\frac{\sin(\pi f a)}{\pi f a}\right)^2$. As a result, there is only a single impulse at $f = 0$. Figure 69 shows the power spectra of a sampled anayltic signal $x(t) = e^{j2\pi f_0 t}$, where $f_0 = 5$ Hz. The number of samples is $N = 1024$. The uniform interval is given by $T = 1/3$ s. The jitters are uniformly distributed in $[-a/2, a/2]$, where $a = T$. The random amplitude process has a correlation coefficient

of $\rho = 0.9$. In the uniform sampling case, an aliasing frequency $f_0 - 2F_s = -1$ Hz is shown in the Nyquist range. In the uniformly distributed JRS case, no aliasing is present and the noise floor is non-flat. Figure 70 shows the signal power spectra similar to the case in Figure 69. The only difference is that $\rho = 1$ in this case. Figure 71 shows the power signal spectra similar to the case in Figure 70. The only difference is that $a = T/2$. Note that aliases are present this time. The aliasing noise floor is also non-flat.

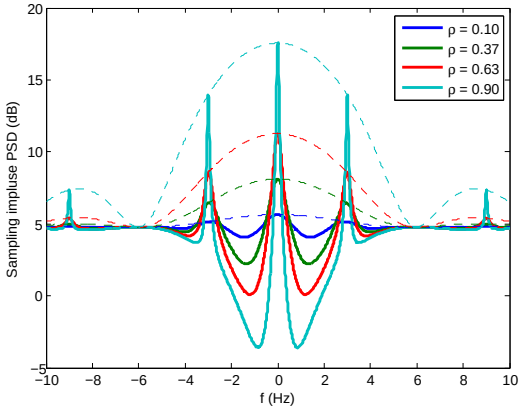

Figure 67: The shape of $\Phi_s(f)$ as a function of ρ for uniformly distributed JRS. $a = T/2$.

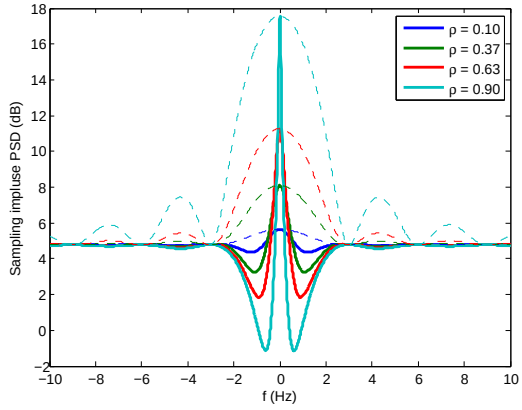

Figure 68: The shape of $\Phi_s(f)$ as a function of ρ for uniformly distributed JRS. $a = T$.

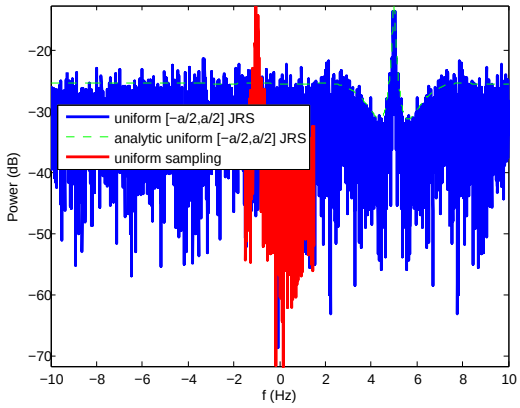

Figure 69: The power spectra of a sampled analytic signal with a frequency at 5 Hz for uniformly distributed JRS. $T = 1/3$ s, $a = T$, $\rho = 0.9$, N = 1024.

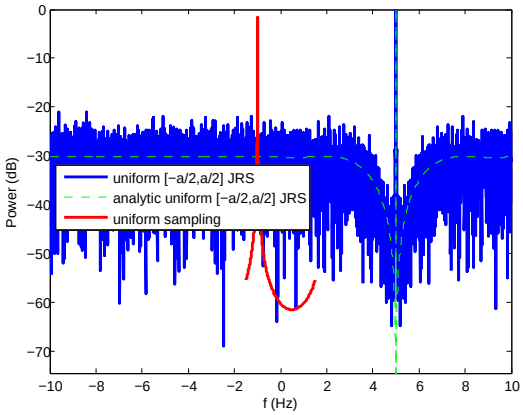

Figure 70: The power spectra of a sampled analytic signal with a frequency at 5 Hz for uniformly distributed JRS. $T = 1/3$ s, $a = T$, $\rho = 1$, N = 1024.

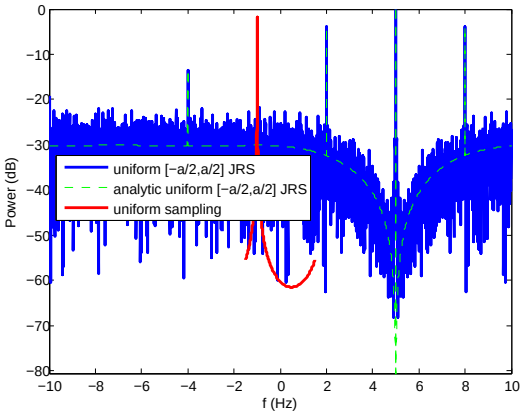

Figure 71: The power spectra of a sampled analytic signal with a frequency at 5 Hz for uniformly distributed JRS. $T = 1/3$ s, $a = T/2$, $\rho = 1$, N = 1024.

3.1.3 Random Skip Sampling (RSS)

Consider the uniform sampling case where

$$f_{\tau_k}(x) = \delta(x - T), \tag{258}$$

$$\psi_{\tau_k}(2\pi f) = e^{j2\pi fT}. \tag{259}$$

when $|\rho| < 1$,

$$\Phi_s(f) = \frac{1}{T} P(\rho, 2\pi fT) \tag{260}$$

$$= \frac{1}{T} \frac{1 - \rho^2}{1 - 2\rho\cos(2\pi fT) + \rho^2}. \tag{261}$$

$\Phi_s(f)$ is periodic in $\frac{1}{T}$.

$$\Phi_s(f) = \Phi_s(f - \frac{n}{T}), \quad \forall n \in \mathbb{Z}. \tag{262}$$

when $\rho = 1$,

$$\Phi_s(f) = \frac{2\pi}{T} \sum_{n=-\infty}^{\infty} \delta(2\pi fT - 2\pi n) = \frac{1}{T^2} \sum_{n=-\infty}^{\infty} \delta(f - \frac{n}{T}). \tag{263}$$

Therefore, the PSD of the uniformly sampled signal is the superposition of a periodic repetitions of the signal's PSD. Aliasing occurs when the original signal is not band-limited to $[-\frac{1}{2T}, \frac{1}{2T}]$. When combining random skip sampling with uniform sampling,

$$\psi_{\tau'_k}(2\pi f) = \frac{(1-q)e^{j2\pi fT}}{1 - qe^{j2\pi fT}}. \tag{264}$$

When $|\rho| < 1$,

$$\Phi_s(f) = \frac{1-q}{T} \Re\left[\frac{1 + \rho \frac{(1-q)e^{j2\pi fT}}{1-qe^{j2\pi fT}}}{1 - \rho \frac{(1-q)e^{j2\pi fT}}{1-qe^{j2\pi fT}}} \right] \tag{265}$$

$$= \frac{1-q}{T} \Re\left[\frac{\rho(1-q)}{\mu} \frac{1 + \mu e^{j2\pi fT}}{1 - \mu e^{j2\pi fT}} + \frac{q}{\mu} \right] \tag{266}$$

$$= \frac{1-q}{T} \left\{ \frac{\rho(1-q)}{\mu} P(\mu, 2\pi fT) + \frac{q}{\mu} \right\}, \tag{267}$$

where

$$\mu = \rho - \rho q + q \in (-1, 1). \tag{268}$$

when $\rho = 1$, we have $\mu = 1$,

$$\Phi_s(f) = \frac{1-q}{T}\left\{\frac{1-q}{T}\sum_{n=-\infty}^{\infty}\delta(f-\frac{n}{T}) + q\right\}. \tag{269}$$

As we have derived, $\Phi_s(f)$ is still periodic in T. Therefore, uniform sampling with random skip cannot eliminate aliasing. The result is quite intuitive. Since uniform sampling without skipping cannot eliminate aliasing, discarding samples randomly cannot eliminate the aliasing effect. But one interesting result here is that the aliasing periodicity T is unrelated to the skip probability q, a higher skip probability only increases the aliasing noise floor. Therefore, we can reduce the average sampling frequency by applying a high skip probability q at the cost of an elevated aliasing noise floor.

Figure 72 shows the power spectra of a sampled analytic signal $x(t) = e^{j2\pi f_0 t}$, where $f_0 = 5$ Hz. The uniform time interval is $T = 1/17$ s or $F_s = 17$ Hz. The skip probability is $q = 0.1$. The number of samples is $N = 1024$. The power spectra of the sampled signal are only shown in the range of $[-F_s/2, F_s/2]$. Figure 73 shows the power spectra similar to the case in Figure 72 except for $q = 0.9$. In both cases, the aliasing noise floor is flat and is proportional to q.

3.1.4 Connections with Compressive Sensing

The Fourier random sampling in compressive sensing (CS) can be well understood using the framework of random skip sampling. In Fourier random sampling, we are performing independent Bernoulli trials on a uniform sampling grid to decide whether to sample or skip the signal at the current time point. If we denote the time interval on a uniform sampling grid as T, the expected value of the sampling frequency, or the average sampling frequency, is $\beta = (1-q)/T$, where q is the skip probability. Suppose the signal $x(t)$ is band-limited in $[-1/2T, 1/2T]$, the random skips sampling

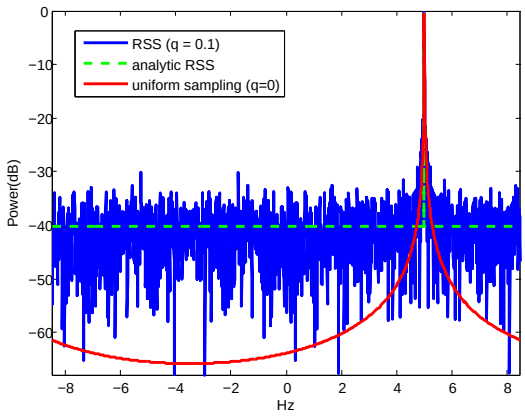

Figure 72: The power spectra of a sampled analytic signal with a frequency at 5 Hz for RSS. $F_s = 17$ Hz, $q = 0.1$, N $= 1024$.

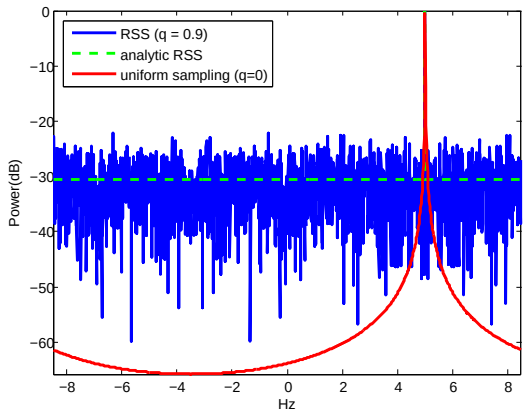

Figure 73: The power spectra of a sampled analytic with a frequency at 5 Hz for RSS. $F_s = 17$ Hz, $q = 0.9$, N $= 1024$.

scheme is able to capture the signal without aliasing in $[-1/2T, 1/2T]$ regardless of how large the skip probability q is. Denote N as the number of uniform time points in the acquisition time window, the expected number of non-uniform samples is

$$M = (1 - q)N. \tag{270}$$

According to the analysis in section 3.1.3, if we sample an analytic signal $x(t) = A_0 e^{j2\pi ft}$ by RSS with skip probability q and obtain M non-uniform samples, there will be a flat aliasing noise floor with power $|A_0|^2 q/M$ in the normalized power spectrum. Now, suppose the input signal is composed of K analytic components $x(t) = \sum_{k=1}^{K} A_k e^{j2\pi f_k t}$, the aliasing noise floor power will increase to $\sum_{k=1}^{K} |A_k|^2 q/M$,

which might bury the weaker frequency components in $x(t)$. To avoid this, we have to make sure that the power of the weakest signal frequency component is C times higher than the power of the aliasing noise floor:

$$\min_k |A_k|^2 = C \sum_{k=1}^{K} |A_k|^2 q/M, \tag{271}$$

where $C > 1$ is a scaling constant. This can be achieved by either decreasing the skip probability q if the acquisition time window is fixed, or increasing the number of samples at the same skip probability. Consider the first case, where the time window is fixed, i.e., N is fixed, and further simplify the model by assuming the magnitude of all frequency components are equal so that

$$q = \frac{M}{CK}. \tag{272}$$

Combining (270) and (272), we have

$$M = CK \frac{N}{N + CK}. \tag{273}$$

The sparsity assumption in compressive sensing requires that $N \gg K$, we can then establish a coarse estimate of M as

$$M \approx CK. \tag{274}$$

Other than a sub-linear $\log N$ term, the number of samples estimated by (274) is the same as the one in compressive sensing. Both estimates imply that the number of samples are mainly proportional to the signal sparsity K. The reconstruction in compressive sensing can also be understood under the RSS framework. If the spectrum of $X(f)$ only contains a small number of frequency components, the impact of the aliasing noise floor on those frequency components is negligible. Therefore, it is possible to reconstruct the signal from the aliasing noise contaminated spectrum.

3.1.5 Time Quantization

All the random sampling schemes except RSS mentioned in Section 3.1 are based upon the assumption that the time intervals $\{\tau_k\}$ and jitters $\{u_k\}$ are continuously distributed. In practice, such a sampling scheme is difficult to implement. The intervals or jitters are usually quantized onto a fixed time grid. For example, consider the case of ARS. Suppose the time grid has a uniform spacing of Δ, denote $\{\tau_k^q\}$ as the quantized time intervals, which are determined according to the nearest neighbor criterion:

$$\tau_k^q = n\Delta \quad \text{if} \quad (n-1/2)\Delta < \tau_k \leq (n+1/2)\Delta \quad \text{for all} \quad n \in \Omega, \tag{275}$$

where Ω is the corresponding feasible integer set. For example, if $\{\tau_k\}$ follows an exponential distribution, $\Omega = \{1, 2, ...\}$. If $\{\tau_k\}$ follows a uniform distribution in $[0, T]$, $\Omega = \{1, 2, ..., \lfloor \frac{T}{\Delta} \rfloor\}$. The PDF of τ_k^q are actually a probability mass function (PMF):

$$f_{\tau_k^q}(x) = \sum_{n\in\Omega} p(n\Delta)\delta(x - n\Delta), \tag{276}$$

where $p(n\Delta)$ stands for the probability of the event $(n-1/2)\Delta < \tau_k \leq (n+1/2)\Delta$ and $\sum_{n\in\Omega} p(n\Delta) = 1$.

$$p(n\Delta) = F_{\tau_k}((n+1/2)\Delta) - F_{\tau_k}((n-1/2)\Delta), \tag{277}$$

where $F_{\tau_k}(x)$ is the cumulative density function (CDF) of τ_k:

$$F_{\tau_k}(x) = \int_{-\infty}^{x} f_{\tau_k}(v)dv. \tag{278}$$

It follows that

$$p(x) = F_{\tau_k}(x + \Delta/2) - F_{\tau_k}(x - \Delta/2). \tag{279}$$

The characteristic function of τ_k^q is given by

$$\psi_{\tau_k^q}(2\pi f) = \int_{-\infty}^{\infty} \sum_{n\in\Omega} p(n\Delta)\delta(x-n\Delta)e^{j2\pi fx}dx \tag{280}$$

$$= \sum_{n\in\Omega} p(n\Delta)e^{j2\pi fn\Delta} \tag{281}$$

$$= \frac{1}{\Delta}\sum_{m=-\infty}^{\infty} P^*(f-\frac{m}{\Delta}), \tag{282}$$

where $P(f) = \mathcal{F}\{p(x)\}$ is the CTFT of $p(x)$, * stands for conjugate. Note that

$$\mathcal{F}\{F_{\tau_k}(x)\} = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{x} f_{\tau_k}(v)dv\right]e^{-j2\pi fx}dx \tag{283}$$

$$= \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} f_{\tau_k}(v)u(x-v)dv\right]e^{-j2\pi fx}dx \tag{284}$$

$$= \int_{-\infty}^{\infty} f_{\tau_k}(x) * u(x)e^{-j2\pi fx}dx \tag{285}$$

$$= \psi^*_{\tau_k}(2\pi f)U(f), \tag{286}$$

where $u(x)$ is a unit step function and $U(f) = \mathcal{F}\{u(x)\} = \frac{1}{2}(\delta(f) - \frac{j}{\pi f})$.

$$P(f) = \mathcal{F}\{F_{\tau_k}(x+\Delta/2)\} - \mathcal{F}\{F_{\tau_k}(x-\Delta/2)\} \tag{287}$$

$$= \psi^*_{\tau_k}(2\pi f)U(f)(e^{j\pi f\Delta} - e^{-j\pi f\Delta}) \tag{288}$$

$$= \psi^*_{\tau_k}(2\pi f)(\delta(f) - \frac{j}{\pi f})j\sin(\pi f\Delta) \tag{289}$$

$$= \Delta\psi^*_{\tau_k}(2\pi f)\frac{\sin(\pi f\Delta)}{\pi f\Delta} \tag{290}$$

Finally, we have

$$\psi_{\tau_k^q}(2\pi f) = \sum_{m=-\infty}^{\infty} \psi_{\tau_k}(2\pi(f-\frac{m}{\Delta}))\frac{\sin(\pi(f\Delta-m))}{\pi(f\Delta-m)} \tag{291}$$

After time quantization, the characteristic function of τ_k^q becomes periodic with a period of $\frac{1}{\Delta}$, i.e.,

$$\psi_{\tau_k^q}(2\pi\frac{m}{\Delta}) = \psi_{\tau_k}(0) = 1. \tag{292}$$

As a result, $\Phi_s(f)$ also becomes periodic. We can ensure that the sampled signal is aliasing free only if $x(t)$ is band-limited in $[-\frac{1}{2\Delta}, \frac{1}{2\Delta}]$. Figure 74 shows the power

spectra of a sampled analytic signal similar to the case in Figure 63. The only difference is that the sampling time is now quantized with a minimal spacing of $\Delta = \frac{1}{5\lambda} = 1/15$ s. As a result the power spectrum of the non-uniformly sampled signal is periodic every 15 Hz. In conclusion, the time quantization granularity Δ determines the highest frequency that can be acquired without aliasing frequencies.

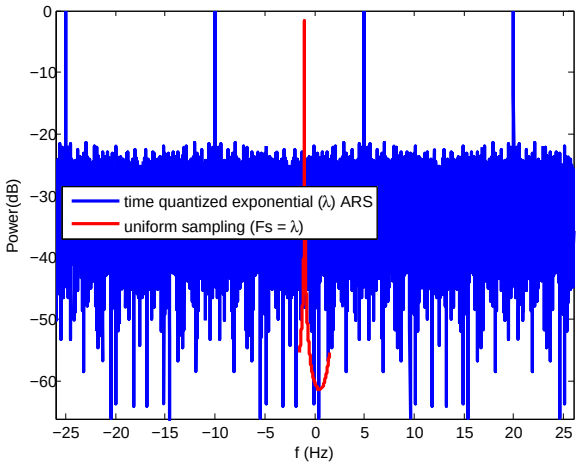

Figure 74: The power spectra of a sampled analytic signal with a frequency at 5 Hz for exponentially distributed ARS. $\lambda = 3$ Hz, $\Delta = \frac{1}{5\lambda}$, $\rho = 1$, N = 1024.

A similar result can be drawn for interval quantization using a round-up criterion:

$$\tau_k^q = n\Delta, \text{ if } (n-1)\Delta < \tau_k \leq n\Delta \tag{293}$$

We can simply the notation of $p(n\Delta)$ to $p[n]$ with the constraint

$$\sum_{n\in\Omega} p[n] = 1. \tag{294}$$

According to the periodicity property of $\tau_k^q(2\pi f)$, the characteristic function of τ_k^q can be re-defined as

$$\psi_{\tau_k^q}(e^{j\omega}) = \sum_{n\in\Omega} p[n]e^{j\omega n}. \tag{295}$$

This definition is similar to the definition of DTFT. As a result, $\psi_{\tau_k^q}(e^{j\omega})$ becomes periodic in 2π. The normalized frequency ω is related to the nominal frequency f according to

$$\omega = 2\pi f\Delta. \tag{296}$$

Accordingly, define the normalized aliasing noise power of $s(t)$ as

$$\Phi_n(e^{j\omega}) = \Re\left\{\frac{1+\psi_{\tau_k^q}(e^{j\omega})}{1-\psi_{\tau_k^q}(e^{j\omega})}\right\}, \quad \omega \in (0,\pi]. \tag{297}$$

It is not difficult to verify that the following results hold:

$$\int_{0^+}^{\pi} \Phi_n(e^{j\omega})d\omega = (1 - \frac{1}{\sum_{n\in\Omega} np[n]})\pi. \tag{298}$$

Note that

$$E[\tau_k^q] = \Delta \sum_{n\in\Omega} np[n], \tag{299}$$

the above equation can be re-written as

$$\frac{1}{\pi}\int_{0^+}^{\pi} \Phi_n(e^{j\omega})d\omega + \frac{\Delta}{E[\tau_k^q]} = 1. \tag{300}$$

Equation (300) represents the fundamental tradeoffs between the sampling frequency and aliasing noise power. We can reduce the average aliasing noise power $\frac{1}{\pi}\int_{0^+}^{\pi} \Phi_n(e^{j\omega})d\omega$ by increasing the normalized average sampling frequency $\frac{\Delta}{E[\tau_k^q]}$ or vice versa. As an extreme case, when $E[\tau_k^q] = \Delta$, which is exactly the uniform sampling case with spacing Δ, the aliasing noise power will decrease to zero.

With the above mentioned definitions, let's re-visit the exponentially distributed ARS with interval quantization of granularity Δ using round-up criterion.

$$\begin{aligned} p[n] =& F_{\tau_k}((n-1)\Delta) - F_{\tau_k}(n\Delta) && (301)\\ =& 1 - e^{-\lambda(n+1)\Delta} - 1 + e^{-\lambda n\Delta} && (302)\\ =& (e^{-\lambda\Delta})^{n-1}(1-e^{-\lambda\Delta}) && (303)\end{aligned}$$

Recall that in random skip sampling with a skip probability of q, the interval follows a geometrical distribution with

$$p[n] = q^{n-1}(1-q) \tag{304}$$

which is equivalent to equation (303) with the the equivalency:

$$q = e^{-\lambda\Delta}. \tag{305}$$

Therefore, exponentially distributed ARS with interval quantization is equivalent to random skip sampling. For random skip sampling,

$$\psi_{\tau_k^q}(e^{j\omega}) = \frac{(1-q)e^{j\omega}}{1-qe^{j\omega}}, \tag{306}$$

$$\Phi_n(e^{j\omega}) = \Re\left\{\frac{1-qe^{j\omega}+(1-q)e^{j\omega}}{1-qe^{j\omega}-(1-q)e^{j\omega}}\right\} \tag{307}$$

$$= \Re\left\{\frac{1-(2q-1)e^{j\omega}}{1-e^{j\omega}}\right\} \tag{308}$$

$$= q \text{ for } \omega \in (0,\pi] \tag{309}$$

The aliasing noise power is consistent with the expression in (269).

$$E[\tau_k^q] = \frac{\Delta}{1-q}. \tag{310}$$

It is easy to verify that equation (300) is satisfied.

3.2 Aliasing Power Shaping

As indicated in equation (300), once the average sampling frequency is fixed, so is the average aliasing noise power. However, we have the flexibility of shaping the distribution of the aliasing power over the bandwidth $(0,\pi]$. This section explores the aliasing noise power shaping effects of various discretely distributed intervals and jitters.

3.2.1 Discrete Uniform ARS

Let's examine the quantized uniform ARS case:

$$p[n] = 1/M, \text{ for } n = 1, ..., M, \tag{311}$$

$$E[\tau_k^q] = \frac{(M+1)M}{2}\Delta, \tag{312}$$

$$\psi_{\tau_k^q}(e^{j\omega}) = \frac{1}{M}\frac{e^{j\omega}-e^{j\omega(M+1)}}{1-e^{j\omega}} \tag{313}$$

$$= \frac{1}{M}\frac{\sin(\omega M/2)}{\sin(\omega/2)}e^{j\omega M/2}, \tag{314}$$

$$\Phi_n(e^{j\omega}) = P(\frac{1}{M}\frac{\sin(\omega M/2)}{\sin(\omega/2)}, \omega M/2), \tag{315}$$

where $P(r,\theta)$ is the Poisson kernel function. Figures 75 through 77 compare the aliasing noise power function $\Phi_n(e^{j\omega})$ for discrete uniform distributed and geometric distributed intervals with the constraint that the average sampling interval $E[\tau_k^q]$ remains the same in each case. As we decrease the average sampling frequency by increasing M. The average aliasing power is reduced. The aliasing noise power exhibits different distributions in the bandwidth $(0,\pi]$. Compared with geometric distributed ARS, which introduces a flat aliasing noise floor, the aliasing noise floor introduced by discrete uniform ARS manifests a dip at $\omega = 0$, which corresponds to the original signal frequency component. This is favorable as the aliasing noise introduced by sampling an analytic signal with a frequency at f_0 using discrete uniform ARS will has less power concentrated around f_0, which makes the power spectrum of the signal obtain a higher SNR at f_0. However, $\Phi_n(e^{j\omega})$ in discrete uniform ARS also shows a degree of overshooting and oscillation in the frequency band $(0,\pi]$. This is unfavorable but unavoidable because once the average aliasing noise power is fixed, a dip at the low frequency band unavoidably results in some overshoots in the high frequency band. A strong overshoot might be mistaken for a weak frequency component in the power spectrum.

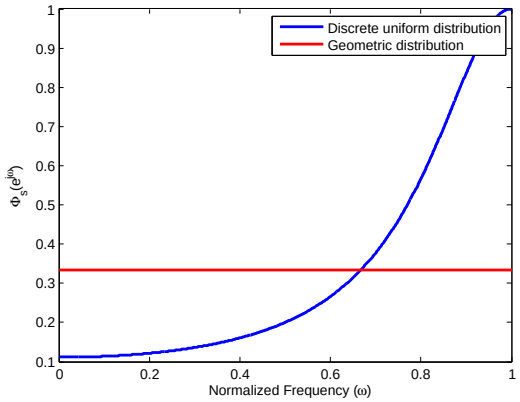

Figure 75: Aliasing noise power function $\Phi_n(e^{j\omega})$ for discrete uniform ARS, M = 2.

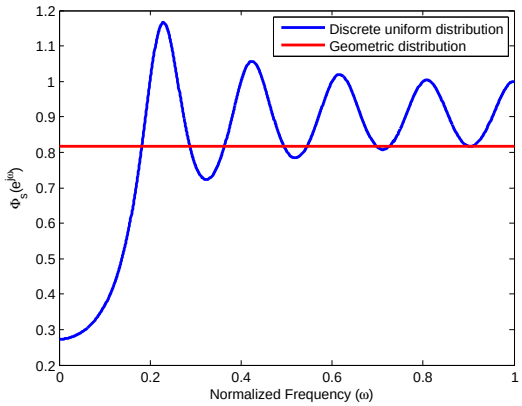

Figure 76: Aliasing noise power function $\Phi_n(e^{j\omega})$ for discrete uniform ARS, M = 10.

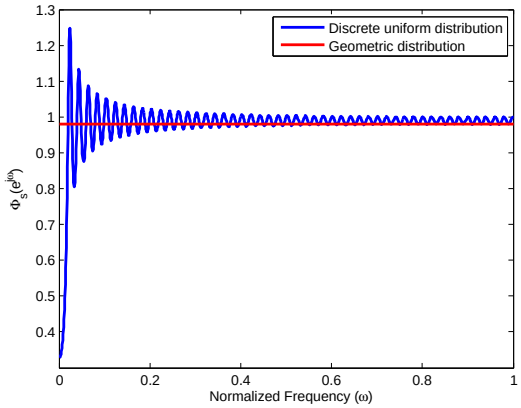

Figure 77: Aliasing noise power function $\Phi_n(e^{j\omega})$ for discrete uniform ARS, M = 100.

For discrete uniform ARS, as indicated in (312), once the average interval or sampling frequency is fixed, so is M. And the aliasing noise power function in (315)is exclusively parameterized by M. It is desirable to introduce a distribution that has more degree of freedom in the design of aliasing noise power function given the average sampling frequency.

3.2.2 Binomial Distributed ARS

The conventional definition of the PMF of a binomially distributed discrete random variable $X \sim \mathcal{B}(N, p)$ is given by

$$p[n] = \binom{N}{n} p^n (1-p)^{N-k}, \quad n = 0, ..., N, \tag{316}$$

which stands for the probability of the n successes in a sequence of N independent Bernoulli trials with success probability of p. In our modified definition, we exclude the zero interval case where $n = 0$ by right shifting the PMF by 1 unit. Accordingly, the characteristic function is multiplied by $e^{j\omega}$. The new PMF is defined as

$$p[n] = \binom{M-1}{n-1} p^{n-1}(1-p)^{M-k}, \quad n = 1, ..., M. \tag{317}$$

$$\psi_{\tau_k^q}(e^{j\omega}) = (1 - p + pe^{j\omega})^{M-1} e^{j\omega}. \tag{318}$$

$$E[\tau_k^q] = ((M-1)p + 1)\Delta. \tag{319}$$

The discrete sampling intervals τ_k^q can be generated as follows. Perform M independent Bernoulli trials with a success probability of p and sum up the number of success trials, denoted as X. $\tau_k^q = (X+1)\Delta$. Given an average sampling interval $E[\tau_k^q]$, Binomial distribution offers the flexibility of tuning the parameters M and p. The minimum feasible M is given by

$$M_{\min} = \begin{cases} \lceil \frac{E[\tau_k^q]}{\Delta} \rceil & \text{if } E[\tau_k^q] \bmod \Delta \neq 0 \\ \lceil \frac{E[\tau_k^q]}{\Delta} \rceil + 1 & \text{if } E[\tau_k^q] \bmod \Delta = 0 \end{cases}. \tag{320}$$

$$p = \frac{E[\tau_k^q]/\Delta - 1}{M - 1}. \tag{321}$$

The aliasing noise power functions $\Phi_n(e^{j\omega})$ for binomially distributed intervals with a fixed average interval length $E[\tau_k^q] = 1.2\Delta$ and different combinations of M and p are shown in Figure 78. The dashed line represents the aliasing noise power introduce by geometric ARS with the same average interval length. Figures 79 and 80 repeat the experiment with $E[\tau_k^q] = 5.2\Delta$ and $E[\tau_k^q] = 10.2\Delta$ respectively. A general observation is that when $E[\tau_k^q]$ is fixed, as M increases, the aliasing noise floor manifests a deeper dip around $\omega = 0$ at the cost of stronger overshoots and fluctuations.

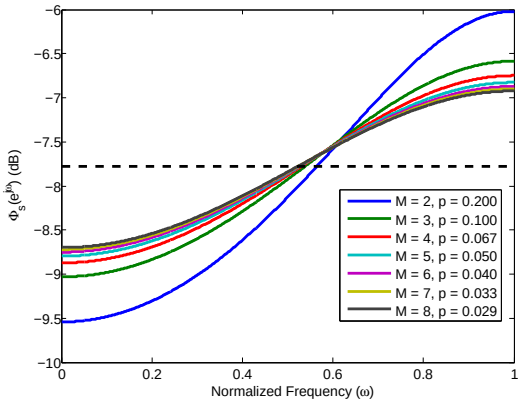

Figure 78: Aliasing noise ower function $\Phi_n(e^{j\omega})$ (dB) for binomial distributed ARS, $E[\tau_k^q] = 1.2\Delta$.

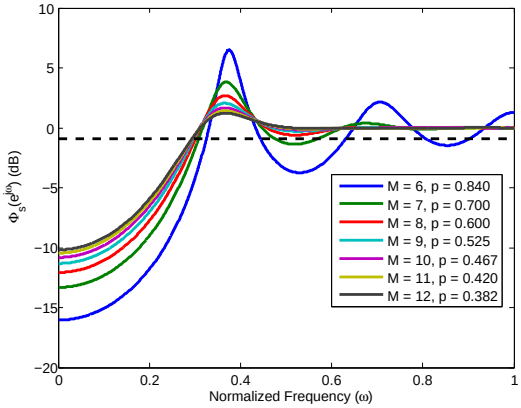

Figure 79: Aliasing noise power function $\Phi_n(e^{j\omega})$ (dB) for binomial distributed ARS, $E[\tau_k^q] = 5.2\Delta$.

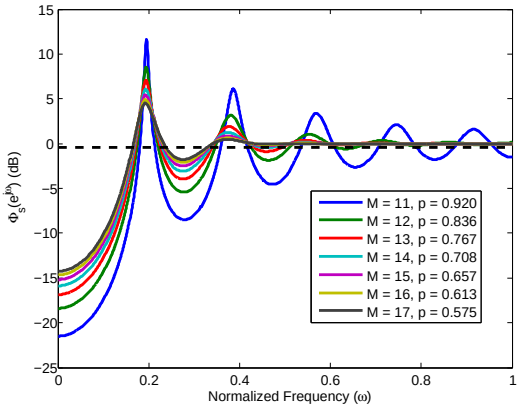

Figure 80: Aliasing noise power function $\Phi_n(e^{j\omega})$ (dB) for binomial distributed ARS, $E[\tau_k^q] = 10.2\Delta$.

3.2.3 Poisson Distributed ARS

A limiting case for the binomial distribution when $M \to \infty$ with its mean $E[\tau_k^q]$ fixed is the Poisson distribution with the following definition:

$$p[n] = \frac{\lambda^{n-1} e^{-\lambda}}{(n-1)!}, \quad n = 1, \ldots \tag{322}$$

$$\psi_{\tau_k^q}(e^{j\omega}) = e^{\lambda(e^{j\omega}-1)} e^{j\omega}. \tag{323}$$

$$E[\tau_k^q] = (\lambda + 1)\Delta. \tag{324}$$

The PMF of three Poisson distributions with $E[\tau_k^q] = 1.2, 5.2, 10.2\Delta$ are shown in Figure 81. The vertical dashed lines stand for the means of the distributions. The corresponding aliasing noise power functions are shown in Figure 82, which are the $M \to \infty$ cases for Figure 78 through 80. The horizontal dashed lines denote the aliasing noise floor of geometric ARS with the same $E[\tau_k^q]$'s.

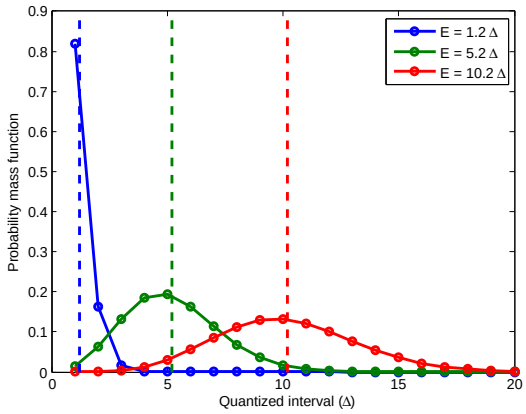

Figure 81: Probability mass function for Poisson distributions.

3.2.4 Negative Binomial Distributed ARS

The conventional definition of the PMF of a negative binomial distributed discrete random variable $X \sim \mathcal{NB}(r, p)$ is given by

$$p[n] = \binom{n+r-1}{n} p^n (1-p)^r, \quad n = 0, \ldots \tag{325}$$

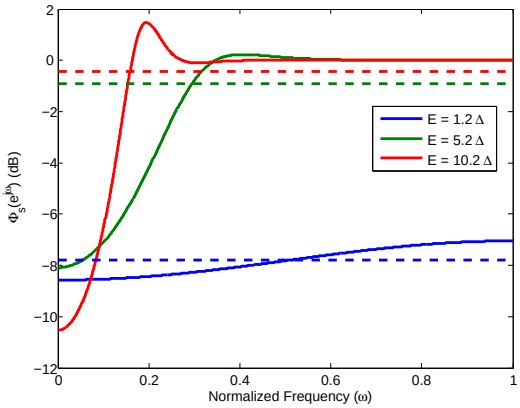

Figure 82: Aliasing noise power function $\Phi_n(e^{j\omega})$ (dB) for Poisson distributed ARS.

which stands for the probability of n successes in a sequence of independent Bernoulli trials until r failures occur with a success probability of p. Obviously, when $r = 1$, negative binomial distribution degenerates into the geometric distribution. With our modified definition,

$$p[n] = \binom{n+r-2}{n-1} p^{n-1}(1-p)^r, \quad n = 1, ... \tag{326}$$

$$\psi_{\tau_k^q}(e^{j\omega}) = \left(\frac{1-p}{1-pe^{j\omega}}\right)^r e^{j\omega}. \tag{327}$$

$$E[\tau_k^q] = (\frac{pr}{1-p} + 1)\Delta. \tag{328}$$

Compared with binomial distributed ARS, negative binomial distributed ARS shows less overshooting and fluctuations in the aliasing noise floor.

3.2.5 Discrete Uniform JRS

We can analyze $\Phi_n(e^{j\omega})$ according to equation (257) for discrete uniform JRS schemes. Suppose the jitters u_k^q are discrete and uniformly distributed in $[-T/2, T/2]$ so that

$$\Pr[u_k^q = -T/2 + mT/M] = 1/M, \quad m = 0, ..., M-1. \tag{329}$$

The time quantization granularity is then

$$\Delta = T/M. \tag{330}$$

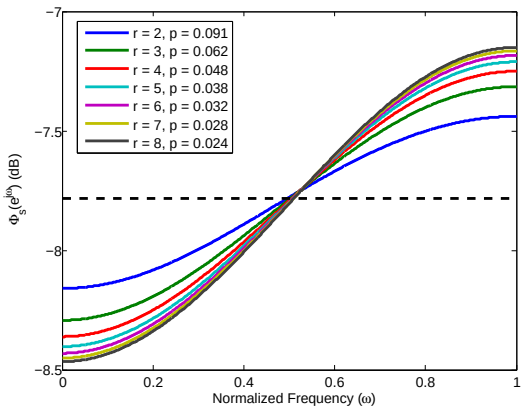

Figure 83: Aliasing noise power function $\Phi_n(e^{j\omega})$ (dB) for negative binomial distributed ARS, $E[\tau_k^q] = 1.2\Delta$.

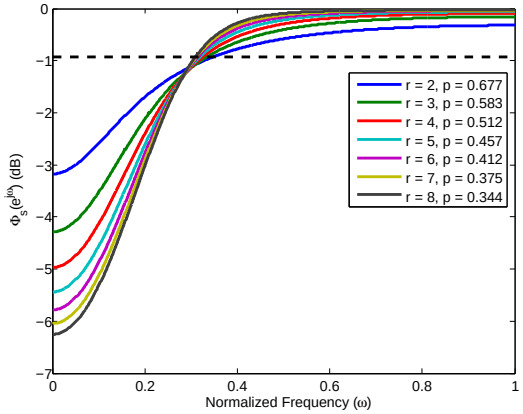

Figure 84: Aliasing noise power function $\Phi_n(e^{j\omega})$ (dB) for negative binomial distributed ARS, $E[\tau_k^q] = 5.2\Delta$.

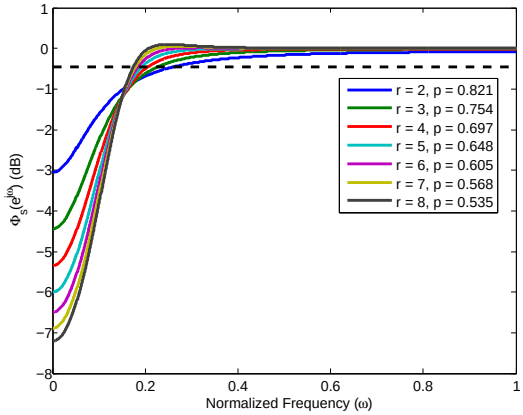

Figure 85: Aliasing noise power function $\Phi_n(e^{j\omega})$ (dB) for negative binomial distributed ARS, $E[\tau_k^q] = 10.2\Delta$.

$$\psi_{u_k^q}(e^{j\omega}) = \frac{1}{M} e^{-\frac{1}{2}j\omega M} \sum_{m=0}^{M-1} e^{j\omega m}. \tag{331}$$

According to equation (257),

$$\Phi_n(e^{j\omega}) = 1 - \frac{1}{M^2} \left(\frac{\sin(\omega M/2)}{\sin(\omega/2)} \right)^2. \tag{332}$$

By increasing M, we can increase the frequency range $[-\frac{M}{2T}, \frac{M}{2T}]$ that is free from aliasing frequencies at a fixed average sampling frequency of $\frac{1}{T}$. A comparison of $\Phi_n(e^{j\omega})$ between discrete uniform JRS and geometric distributed ARS is shown in Figure 86. The time quantization granularity Δ and average sampling interval are kept the same in both sampling schemes. We have

$$q = 1 - \frac{1}{M}. \tag{333}$$

The dashed lines in Figure 86 represent the aliasing noise power with the corresponding geometric distributed ARS. A remarkable feature of the discrete uniform JRS scheme is that

$$\Phi_n(e^{j0}) = 0, \tag{334}$$

$$\Phi_n(e^{j\pi}) \leq 1, \tag{335}$$

which means the aliasing noise floor has no impact at the signal frequency. This makes it easier to extract the signal frequencies from the aliasing noise floor.

3.3 Random Sampling Architectures and Reconstruction Algorithms

This section is dedicated to applying the random sampling theory introduced in the previous sections to obtain practical implementations of random sampling on standard ADC architectures. First, two generic random sampling architectures are proposed to implement discretely distributed ARS and JRS. A new reconstruction algorithm called successive sine matching pursuit is then introduced to reconstruct the signal onto a uniform time grid. Finally, three detailed implementations of random sampling on the SAR ADC, the ramp ADC and the level crossing ADC are proposed and analyzed.

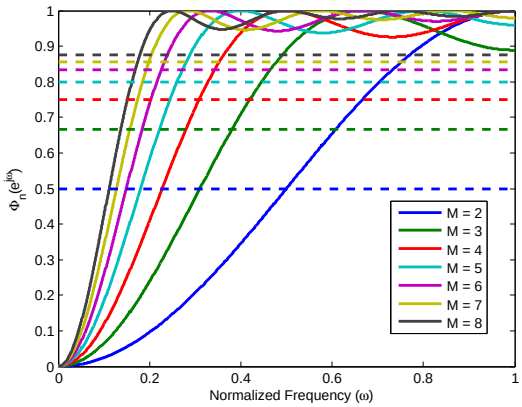

Figure 86: Aliasing noise power $\Phi_n(f)$ for discrete uniform JRS.

3.3.1 Generic Random Sampling Architectures

The random Fourier sampling matrix in compressive sensing or RSS is just one random sampling scheme. Other schemes such as ARS and JRS are also shown to be able to suppress aliasing. Therefore, both ARS and JRS are compatible with compressive sensing theory. Figure 87 shows the block diagram of a generic ARS scheme. The random number generator produces a discrete interval $\tau_k^q = n_k\Delta$ according to a prescribed PDF. The input signal $x(t)$ is sampled at t_k and quantized to $x^q(t_k)$. Finally $x^q(t_k)$ and the interval integer n_k are encoded together to produce a single sample package. A similar generic JRS architecture is shown in Figure 88.

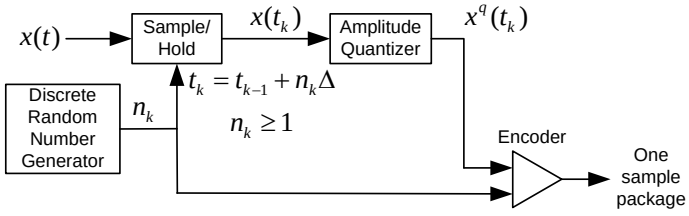

Figure 87: The block diagram of a generic ARS scheme.

If the amplitude of each sample $x(t_k)$ can be quantized to full precision within the minimal interval Δ, then the amplitude quantization noise can still be treated as uniformly distributed additive noise. However, a fundamental tradeoff in almost all ADC architectures is the tradeoff between amplitude resolution and sampling speed.

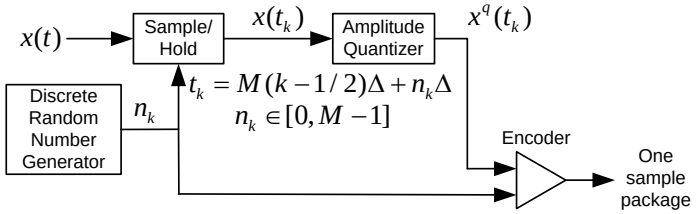

Figure 88: The block diagram of a generic JRS scheme.

Allowing each sample to reach its full amplitude precision increases the minimal time interval Δ and reduces the frequency coverage.

3.3.2 Sparse Multi-sine Signal Recovery via Successive Sine Matching Pursuit

As shown in Section 3.1.4, not all signals are perfectly reconstructable under random sampling. This section concentrate on the reconstruction of a class of spectrally sparse signals. Multi-sine signals are widely used to provide a periodic, well-characterized waveform that can simulate a large variety of complex modulated radio frequency signals in wireless telecommunications. For example, multi-sines are used for system testing [72, 60] and modeling [59, 85], technology standard verification [29] and calibration [84].

3.3.2.1 Power Spectrum

If we are only interested in the detection of certain frequency components from the non-uniform samples, calculating the power spectrum is sufficient. The time quantization in the ARS architecture makes it possible to calculate the power spectrum using an FFT by replacing missing values with zeros. After we have collected M quantized samples with $\{x^q(t_k), n_k\}, k = 1, ..., M$ and $N = \sum_{k=1}^{M} n_k$, we can insert zeros in between each sample according the n_k as shown in Figure 89. If we denote the minimal time interval as Δ and the zero-inserted signal vector as $\bar{\mathbf{x}}$, then we can calculate the normalized power spectrum an N-point FFT

$$\mathbf{p} = \frac{1}{M^2}|FFT\{\bar{\mathbf{x}}\}|^2, \tag{336}$$

where the frequency grid spacing is $\frac{1}{N\Delta}$ Hz. We can denote this kind of interval zero insertion operation as "IZI". In contrast, we denote the conventional trailing zero padding operation as "TZP".

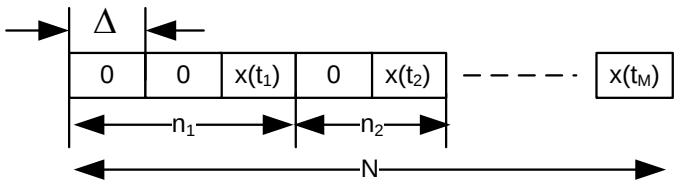

Figure 89: Interval zero insertion for ARS samples.

3.3.2.2 Spectral Leakage and Noise Floor

To evaluate the ARS scheme we consider a classic problem in compressive sensing, where the input signal is assumed to be a multi-sine signal with a small number of frequency components. According to ARS, we obtain M non-uniform quantized samples $\{x^q(t_k)\}$ and the normalized power spectrum can be evaluated according to (336). However, there will be a spectral leakage effect in the form of multiple sidelobes around each frequency bin due to the finite length acquisition time window. In addition, as derived for the ARS, JRS, RSS schemes, there will be an aliasing noise floor in the power spectrum. Also, the amplitude quantization introduces another source of noise that could reshape the total noise floor. All these factors mean that the signal is not perfectly sparse in the frequency domain. Even worse, a strong sidelobe may also over-shadow weaker frequency components in the multi-sine signal. Generic reconstruction algorithms in compressive sensing must be customized to deal with this specific problem.

3.3.2.3 Successive Sine Matching Pursuit

Inspired by the CoSaMP algorithm [55] and the least squares periodogram [73], a new reconstruction algorithm, called successive sine matching pursuit (SSMP), is proposed in this section. The fundamental step in CoSaMP is to iteratively subtract an s sparse

solution, where s is the assumed sparsity of the solution, from the measurement vector so that the power in the signal residual $\mathbf{v}$ can be reduced with each iteration. The SSMP reconstruction algorithm is summarized in the pseudo code of Algorithm 3.

Input: noisy non-uniform samples $\{x^q(t_k), n_k\}, k = 1, ...M$, $N = \sum_{k=1}^{M} n_k$, time quantization granularity Δ, FFT size $N_{FFT} \geq N$, bandwidth search ratio $r \in (0, 1)$.

Output: A reconstructed signal $x[n] = x(n\Delta), n = 0, ..., N-1$

Initialization:

$\Delta f = \frac{r}{N\Delta}$; /* determine the frequency search range */

$\mathbf{x}^0 \in \mathbb{R}^N = \mathbf{0}$; /* initial solution */

$\mathbf{v}^0 \in \mathbb{R}^M = \mathbf{x}_q$; /* initial residual */

$e^0 = \|\mathbf{v}^0\|_2^2$; /* initial residual power */

$k = 0$

Iteration:

repeat

$k = k + 1$

$\bar{\mathbf{v}} = TZP(IZI(\mathbf{v}^{k-1}))$; /* zero insertion and padding */ /* $IZI(\mathbf{v}) \in \mathbb{R}^N$, $\bar{\mathbf{v}} \in \mathbb{R}^{N_{FFT}}$ */ /* frequency sampling spacing = $\frac{1}{N_{FFT}\Delta}$ */

$\mathbf{p} = FFT(\bar{\mathbf{v}})$; /* perform FFT */

$f = |\mathbf{p}|_{(1)}$; /* identify the largest peak in the spectrum */

$f_j \in [f - \Delta f, f + \Delta f]$; /* discretize a frequency band centered around the peak frequency into J frequencies */

$\min_{\alpha_j, \phi_j} = \sum_{k=1}^{M} (v(t_k) - \sum_{j=1}^{J} \alpha_j \cos(2\pi f_j t_k + \phi_j))^2$; /* fit a class of cosine waveforms at f_j to $\mathbf{v}^{k-1}$ */

$v^k(t_k) = v^{k-1}(t_k) - \sum_{j=1}^{J} \alpha_j \cos(2\pi f_j t_k + \phi_j)$; /* excise the cosine waveforms from the residual */

$x^k[n] = x^{k-1}[n] + \sum_{j=1}^{J} \alpha_j \cos(2\pi f_j n\Delta + \phi_j)$; /* add the cosine waveforms to the solution */

$e^k = \|\mathbf{v}^k\|_2^2$; /* calculate the new residual power */

until $e^k \geq e^{k-1}$;

Algorithm 3: SSMP algorithm

Given M non-uniform samples $x^q(t_k), n_k$ with a time quantization granularity of Δ, we can calculate the total number N of uniform time points available in the

acquisition time window. The mainlobe width of each frequency bin is $\frac{2}{N\Delta}$. In the initialization stage, the reconstructed signal $\mathbf{x}$ is initialized as a zero vector and the residual signal $\mathbf{v}$ is initialized as the sampled signal $\mathbf{x}_q$. For each iteration, we first perform interval zero insertion and trail zero padding to $\mathbf{v}$ and perform an FFT on the result. Then the frequency f that corresponds to the largest peak in the spectrum is identified. Since there is amplitude quantization and aliasing noise, the actual signal frequency might deviate slightly from f. Instead of fitting a single sinusoid at exact f, a cluster of sinusoids with frequencies centered around f are used to fit the residual vector. The frequency search range Δf is set to be half the width of the mainlobe scaled by a factor $r \in (0, 1)$. The number of sinusoids is denoted as J. The next step is to fit this cluster of sinusoids at frequencies f_j to the residual vector according to the least squares criterion. Finally, the identified sinusoids are subtracted from the residual vector. At the same time, they are also used to reconstruct the signal on a uniform time grid. The algorithm ends when the power in the residual vector can no longer be reduced.

The advantage of the proposed SSMP algorithm can be best illustrated in the following example. Consider a continuous time signal $x(t)$ that is composed of two sinusoids with amplitudes 1 and 0.05 and frequencies 5.5 and 5.51 Hz. The maximum sampling frequency is $F_s = \frac{1}{\Delta} = 25$ Hz. We are sampling the signal using RSS with a skip probability of $q = 0.9$. Therefore, the average sampling frequency is $0.1F_s$, which is well below the Nyquist frequency of the signal. The number of samples acquired is $M = 1024$ out of the total possible number of uniform time points in the acquisition time window which is $N = \sum_{k=1}^{M} n_k = 10636$. The sampling times are given by $t_k = \sum_{i=1}^{k} \tau_i^q$. The amplitude quantized measurements are given by $x^q(t_k) = x(t_k) + v[k]$, where $v[k]$ is Gaussian white noise with $\sigma = 0.1$. The objective of the reconstruction algorithm is to recover $x[n] = x(n\Delta), n = 0, ..., N-1$ from $x^q(t_k), k = 1, ..., M$.

The first step for reconstruction is to insert zeros in between the signal samples $\{x^q(t_k)\}$ as shown in the previous section. Depending on the frequency sample spacing we want to achieve and the number of samples we have, we can also pad zeros to the end of the IZI vector and perform a longer FFT afterwards. In this example, the FFT size is given by $N_{FFT} = 2^{20}$.

The next step is to identify the largest peak in the spectrum. By contrast, CoSaMP algorithm tries to identify the $2s$ largest components in the spectrum simultaneously, in which case the sidelobes near the strongest main lobe around 5.5 Hz might be picked. CoSaMP then tries to find a solution by minimizing the least squares error on the $2s$ picked frequencies. The issue is that we know the sidelobes are caused by the windowing effect. There are actually no such frequency components in the multi-sine signal. Trying to enforce a least squares solution on those frequencies is likely to introduce many frequency mismatch errors. SSMP avoids this problem by only searching for the strongest peak at each iteration. After identifying the strongest frequency component at f, SSMP tries to fit a cluster of sinusoids within a close vicinity to f to residual signal in a minimum least squares error sense.

Finally, we can subtract the identified sinusoids from the residual signal and add it to the reconstructed signal $x[n]$ at the same time. The above process is repeated until the power in the residual signal cannot be further reduced. The substraction step is very critical because it effectively removes the sidelobes introduced by the identified strongest frequency component as well as reducing the corresponding aliasing noise floor associated with it. As a result, the weaker frequency components should become more detectable in the residual power spectrum.

Figure 90 shows the residual and reconstructed power spectra after the first iteration of SSMP. The frequency component at 5.5 Hz has been identified and removed from the samples. The residual spectrum shows a much stronger peak at 5.51 Hz.

Denote the amplitudes of the sines in the multi-sine signal as $A_k, k = 1, ..., P$,

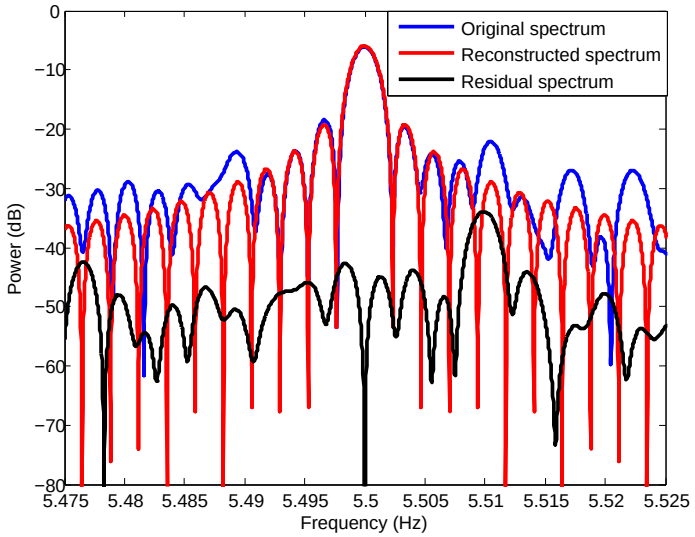

Figure 90: The power spectra after the first iteration of SSMP.

where P is the number of sines. We can index the sine components according to their amplitudes: $A_k \geq A_{k+1}$. The condition of a successful detection of component k is

$$\frac{A_k^2}{4} \gg \sum_{p=k}^{P} A_p^2 \frac{q}{2M} + \frac{\sigma^2}{M}, \tag{337}$$

where $\sum_{p=k}^{P} A_p^2 \frac{q}{2M}$ is the aliasing noise power introduced by RSS and $\frac{\sigma^2}{M}$ is the power of the amplitude quantization noise. In this example, the first sinusoidal component has a power that is -6 dB, which is much larger than the total noise power of -33 dB. After a successful identification and removal of the first sinusoid, the total noise power is reduced to -50 dB, which is still well below the power of the second sinusoidal component, which is -32 dB.

3.3.3 Random Sampling on SAR ADC Architecture

The successive approximation ADC [39] is a type of analog to digital converter that converts analog waveforms using feedback and successive comparisons. Instead of counting up in a binary sequence as in a ramp ADC, a special counter circuit called a successive approximation register (SAR) enables by a binary search through all possible quantization levels starting from the most significant bit (MSB) to the least

significant bit (LSB).

As shown in Figure 91, the analog to digital conversion process starts with an enable signal that commands the sample and hold (S/H) circuit to keep the signal value constant during the conversion cycle. A comparator determines if the S/H input is above or below the DAC output and stores the result in the SAR from MSB to LSB. The timing control determines when the conversion begins and completes.

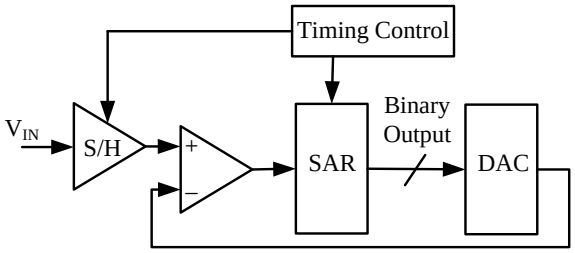

Figure 91: Block diagram of the SAR ADC architecture.

Figure 92 shows the approximation process of a SAR ADC. The SAR is initialized to 1/2 V_{REF}. The comparator monitors whether the output of the DAC is greater than (set the MSB to 0) or less than (set the MSB to 1) the analog signal. The SAR is then set to either 1/4 V_{REF} (if the MSB was set 0) or 3/4 V_{REF} (if the MSB was set 1), so that the comparator can make the decision for the next bit of the conversion. The strategy continues until all bit values have been determined. A fundamental characteristic of the SAR ADC is the tradeoff between sampling speed and quantization level. The SAR ADC can be configured to sample every Δ seconds with 1-bit quantization, or every $K\Delta$ seconds with K-bit quantization. If we set the minimal interval time in Figure 87 or 88 as the single bit approximation time in the SAR ADC, we are maximizing the frequency coverage to $[-\frac{1}{2\Delta}, \frac{1}{2\Delta}]$. However, we pay the price of randomized amplitude quantization over all the samples.

Given the binary output of the SAR ADC, the corresponding voltage can be expressed as $\frac{\text{code}}{2^K}V_{\text{REF}}$, where code is the the ADC output code converted to decimal form. The input range must lie between zero and the reference voltage V_{REF},

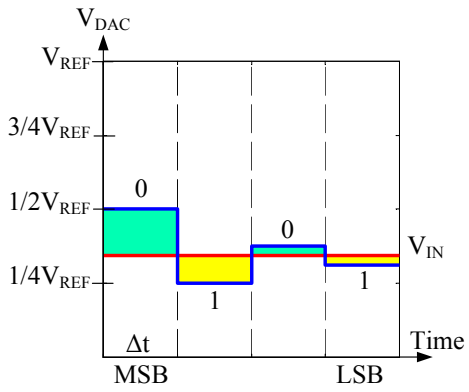

Figure 92: Approximation process of a SAR ADC.

e.g., 2.5 V. There are many industrial applications where a bipolar signal (e.g. $\pm 5V$, $\pm 10V$) needs to be converted. In such cases, some external circuitry using an operational amplifier or resistor network can be employed to perform the attenuation and level shifting needed to match the input range of the SAR ADC. Ideally, the transformation can be expressed as a linear relationship:

$$V_{\text{IN}} = aV_{\text{AIN}} + b \tag{338}$$

where V_{IN} is the final analog signal input to the ADC, V_{AIN} is the analog signal input to the external circuitry, and $\{a, b\}$ are the scaling and shifting parameters. Suppose a sampling interval is given by $k\Delta$, the quantization error of the corresponding sample is given by

$$e_k = V_{\text{IN}} - V_{\text{IN}}^q, \tag{339}$$

where V_{IN}^q is the quantized voltage. Denote the maximum quantization level as K,

$$e_k \sim \begin{cases} U[0, \frac{1}{2^k}]V_{\text{REF}} & k = 1, ..., K-1 \\ U[0, \frac{1}{2^K}]V_{\text{REF}} & k \geq K \end{cases}. \tag{340}$$

Eventually, we need to convert the V_{IN}^q back to the bipolar signal according to

$$V_{\text{AIN}}^q = \frac{1}{a}(V_{\text{IN}}^q - b). \tag{341}$$

The scaled quantization noise is defined by

$$e'_k = V_{\text{AIN}} - V^q_{\text{AIN}} = \frac{1}{a} e_k. \tag{342}$$

Suppose we adopt the RSS scheme with a skip probability of q, then

$$e'_k \sim \begin{cases} U[0, \frac{1}{2^k}]V_{\text{REF}}/a & \text{with probability } p_k = q^{k-1}(1-q) \quad k = 1, ..., K-1 \\ U[0, \frac{1}{2^K}]V_{\text{REF}}/a & \text{with probability } p_K = q^{K-1} \end{cases}. \tag{343}$$

$$E[e'_k] = V_{\text{REF}}/a \sum_{k=1}^{K} \frac{p_k}{2^{k+1}}. \tag{344}$$

The PDFs of the quantization error e'_k for different q's are shown in Figure 93. The values of e'_k are normalized by V_{REF}/a, the PDFs are normalized by a/V_{REF}. The dashed lines denote the means of e'_k. The maximum quantization level is set as $K = 4$. A general observation is that the mean and variance of e'_k decreases as q approaches 1. For q close to 1, we can model the e'_k as uniform distributed white noise with a flat spectrum. To make the mean of the scaled quantization noise e'_k zero, we need to modify equation (341) to

$$V^q_{\text{AIN}} = \frac{1}{a}(V^q_{\text{IN}} - b) + E[e'_k]. \tag{345}$$

In order to reduce the quantization noise power, we have to reduce the variance of e'_k by implying a high skip probability q.

A SAR ADC based random sampling example is shown in Figure 94, where the input signal $x(t)$ is composed of two sinusoids with amplitudes 1 V and 0.1 V at frequencies 5 and 80 Hz. The maximum sampling frequency is $F_s = \frac{1}{\Delta} = 250$ Hz. The skip probability is set as $q = 0.95$ so that the expected average sampling frequency is 1.25 Hz. The number of non-uniform samples is $M = 1024$. $V_{\text{REF}} = 2$ V, so the scaling and bias parameters are given by $a = 1$, $b = 1$ V. The maximum number of amplitude quantization levels is given by $K = 16$.

After sampling M non-uniform samples, we obtain a time window that contains $N = \sum_{k=1}^{M} n_k = 20830$ uniform time points with a total length of 83.32 seconds. The actual average sampling frequency is 12.29 Hz. The amplitude and time quantization bit rate of the sampling process are 136.88 and 49.08 bits/second, respectively.

The spectra of the original signal, the scaled amplitude quantization noise, the recovered signal after 2 iterations of SSMP and the residual signal are shown with the same scale in Figure 94. The flat aliasing noise buries the weak frequency component at 80 Hz in the original spectrum. The scaled quantization noise has a flat spectrum underneath the aliasing noise power spectrum. SSMP successfully recovers the two sinusoids after two iterations, leaving a residual spectrum that approximates the scaled quantization noise spectrum. The signal to quantization noise ration (SQNR) of the recovered signal is 28.72 dB, which corresponds to an effective number of bit (ENOB) of 4.77 bits according to the following relationship

$$\text{ENOB} = \frac{\text{SQNR} - 1.76}{6.02}. \tag{346}$$

Since the signal is recovered on the uniform time grid with frequency F_s, the recovered signal has an equivalent bit rate of 1192.59 bits/second.

In conclusion, SAR ADC based random sampling with an average sampling frequency of 12.29 Hz and a total bit rate of 185.96 bits/second is approximately equivalent to a uniform sampling scheme with a uniform sampling frequency of 250 Hz and a bit rate of 1192.59 bits/second in this example. Therefore, we can sample the same signal with the proposed random sampling scheme at a bit rate 6 times lower than uniform sampling.

3.3.4 Random Sampling on Ramp ADC Architecture

The principle of a ramp ADC is to connect the output of a binary counter to the input of a DAC, then compare the analog output of the DAC with the analog input signal to be digitized and use a comparator to determine when the conversion has

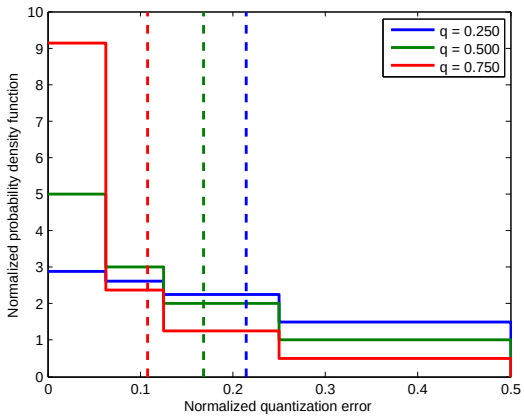

Figure 93: The normalized PDF of the quantization error in SAR ADCs. K = 4.

completed. As shown in Figure 95, the enabling terminal is connected to the output of a comparator. The counter keeps increasing its binary output according to a clock signal until the DAC output is higher than the input analog signal. The counter is then reset to 0 to begin the conversion of the next sample. Figure 96 shows the approximation process of a ramp ADC, where Δ_t and Δ_a are the time and amplitude quantization step size. The conversion time of a sample is proportional to the amplitude of that sample. Therefore, the sampling interval distribution is dependent upon the input signal, but the amplitude quantization is uniform in this case. To further randomize the sampling intervals, we can add an amplifier to the output of the DAC so that we can control the slope of the ramp during the conversion process. As a result, the sampling interval would depend on both the signal amplitude and the gains in the amplifiers. However, a variable slope means that the quantization noise is no longer uniform, which can be seen in Figure fig:rampRecoveredSpectrumSubplots(b).

A ramp ADC based random sampling example is shown in Figure 97, where the input signal $x(t)$ is composed of two sinusoids with amplitudes 1 V and 0.1 V and frequencies 5 and 80 Hz. The maximum sampling frequency is $F_s = \frac{1}{\Delta_t} = 250$ Hz. For each sample, the gain α in the amplifier is set to be discretely and uniformly distributed in $[1, 16]$. The number of non-uniform samples is $M = 1024$. $V_{\text{REF}} = 2$ V. The full amplitude quantization level is $K = 8$, which gives an amplitude step size of

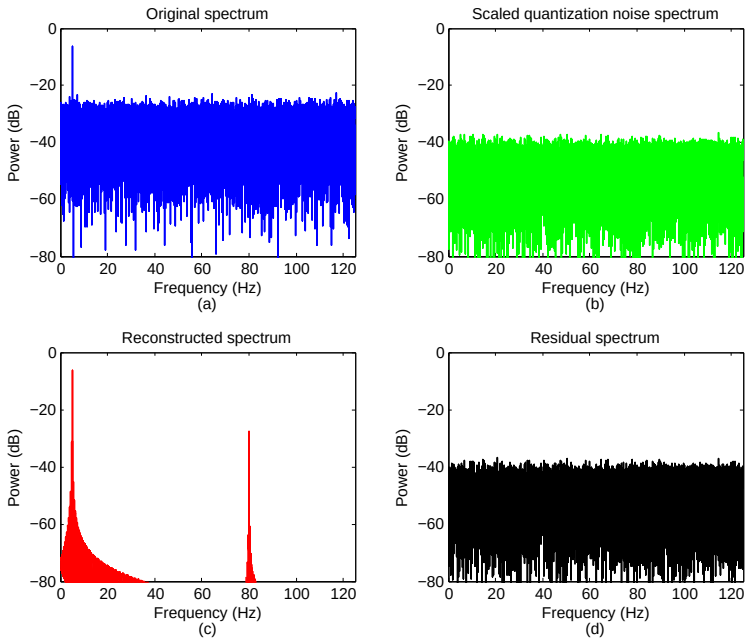

Figure 94: Spectra comparison for SAR ADC based random sampling. (a) The power spectrum of the non-uniformly sampled and quantized signal (b) The power spectrum of the quantization noise (c) The power spectrum of the reconstructed signal though the SSMP algorithm (d) The power spectrum of the residual signal

$\Delta_a = \frac{V_{\text{REF}}}{2^K}$. The scaling and bias parameters are given by $a = 1$, $b = 1$ V.

After sampling M non-uniform samples, we obtain a time window that contains $N = \sum_{k=1}^{M} n_k = 6675$ uniform time points with a total length of 26.7 seconds. The average sampling frequency is 38.35 Hz. The amplitude and time quantization bit rate are 115.81 and 63.11 bits/second, respectively.

The spectra of the original signal, the scaled amplitude quantization noise, the recovered signal after 2 iterations of SSMP and the residual signal are shown with

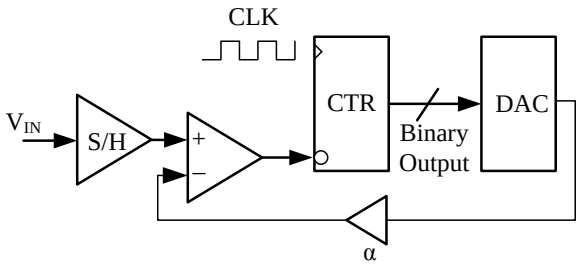

Figure 95: Block diagram of the ramp ADC architecture.

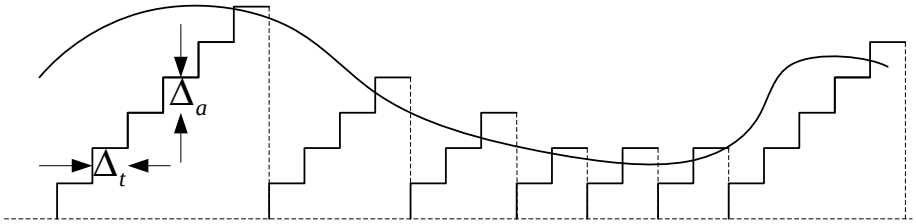

Figure 96: Approximation process of a ramp ADC.

the same scale in Figure 97. Since the sampling interval depends partially on the input signal. Both the aliasing and scaled amplitude quantization noises are correlated with the input signal, which explains the presence of aliasing frequencies in the sampled spectrum. However, SSMP still successfully detects and recovers the two sinusoids after two iterations. The SQNR of the recovered signal is 18.35 dB, which corresponds to an ENOB of 3.05 bits. The recovered signal has an equivalent bit rate of 762.02 bits/second. In conclusion, ramp ADC based random sampling with an average sampling frequency of 38.35 Hz and a total bit rate of 178.92 bits/second is approximately equivalent to a uniform sampling scheme with a uniform sampling frequency of 250 Hz and a bit rate of 762.02 bits/second in this example.

3.3.5 Random Sampling on Level Crossing ADC Architecture

The principle of a level crossing (LC) ADC is that the input signal is sampled whenever it crosses one of a set of preset voltage levels V_1 to V_K. As shown in the logic diagram in Figure 98, the XOR block performs bit-wise XOR operation on the outputs of the comparators and the register. If the comparators' outputs differ from the value stored in the register, one voltage level is crossed by the input signal. A decoder determines which voltage level is crossed according to the outputs of the XOR block and the comparators. Then the counter is reset and the register is enabled to store the current comparator outputs. A waveform sampling example of a level crossing ADC is shown in Figure 99. During the interval t_1, the output vector of the comparators connected with V_1 and V_2 is $[0, 1]$. At the start of interval t_2, the output vector is

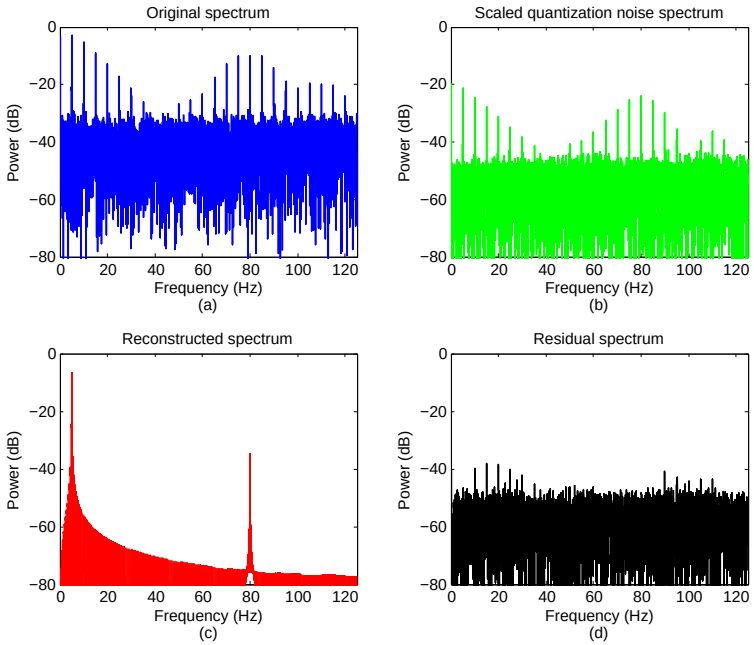

Figure 97: Spectra comparison for ramp ADC based random sampling. (a) The power spectrum of the non-uniformly sampled and quantized signal (b) The power spectrum of the quantization noise (c) The power spectrum of the reconstructed signal though the SSMP algorithm (d) The power spectrum of the residual signal

changed to $[1, 1]$ which indicates voltage level V_1 is crossed. Similarly, at the start of interval t_3, the output vector is updated to $[0, 1]$, which suggests that V_1 is crossed again.

A level crossing ADC based random sampling example is shown in Figure 100, where the input signal $x(t)$ is composed of two sinusoids with amplitudes 1 V and 0.1 V and frequencies 5 and 80 Hz. The maximum sampling frequency, or the clock frequency of the counter, is $F_s = \frac{1}{\Delta} = 250$ Hz. There are $K = 8$ reference voltage levels uniformly distributed in $[-V_{\text{REF}}, V_{\text{REF}}]$, where $V_{\text{REF}} = 1.1$ V. The number of non-uniform samples is $M = 1024$. After sampling M non-uniform samples, we obtain a time window that contains $N = \sum_{k=1}^{M} n_k = 2141$ uniform time points with a total length of 8.56 seconds. The average sampling frequency is 119.57 Hz. The amplitude and time quantization bit rate are 268.57 and 45.07 bits/second, respectively.

The spectra of the original signal, the scaled amplitude quantization error, the

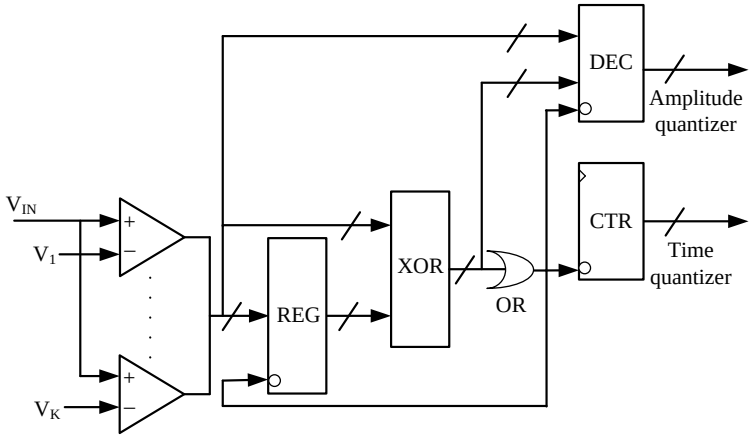

Figure 98: Block diagram of the level crossing ADC architecture.

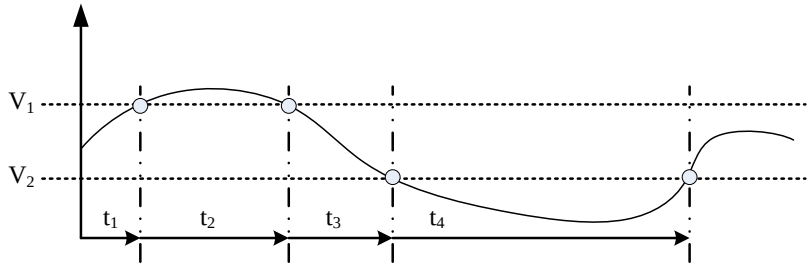

Figure 99: Approximation process of a level crossing ADC.

recovered signal after two iterations of SSMP and the residual signal are shown with the same scale in Figure 97. Similar to the ramp ADC, the sampling interval depends on the input signal and the reference voltage levels. Both the aliasing and amplitude quantization noise are correlated with the input signal, which explains the presence of aliasing frequencies in the sampled spectrum. However, SSMP still successfully detects and recovers the two sine components after two iterations. The SQNR of the recovered signal is 19.11 dB, which corresponds to an ENOB of 3.17 bits. The recovered signal has an equivalent bit rate of 793.68 bits/second. In conclusion, level crossing ADC based random sampling with an average sampling frequency of 119.57 Hz and a total bit rate of 313.64 bits/second is approximately equivalent to a uniform sampling scheme with a uniform sampling frequency of 250 Hz and a bit rate of 793.68 bits/second in this example.

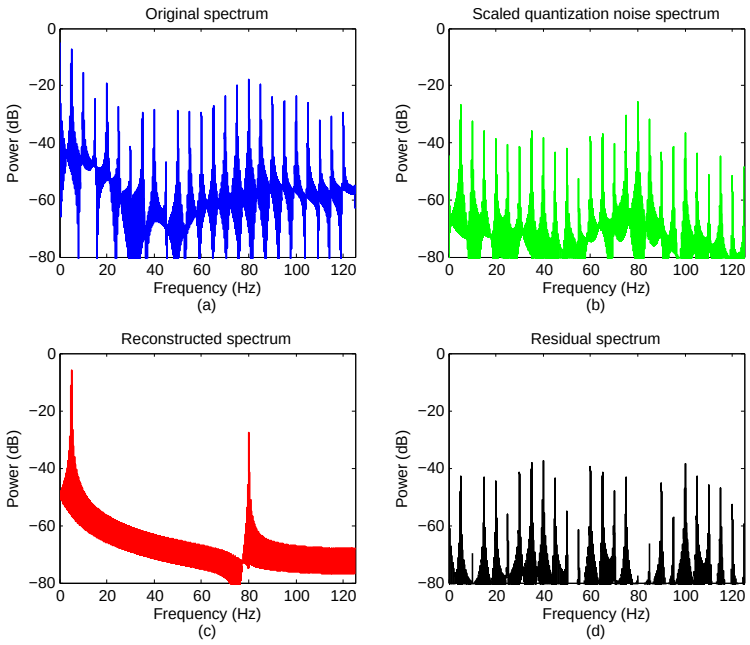

Figure 100: Spectra comparison for level crossing ADC based random sampling. (a) The power spectrum of the non-uniformly sampled and quantized signal (b) The power spectrum of the quantization noise (c) The power spectrum of the reconstructed signal though the SSMP algorithm (d) The power spectrum of the residual signal

A detailed comparison between the performance of the proposed random sampling architectures are listed in Table 7. The fundamental performance metric for any ADC is the bit rate. In uniform sampling, an ADC can be configured either to sample at a low rate (samples/sec.) with a high amplitude resolution (bits/sample) or to sample at a high rate with a low amplitude resolution. In both cases, the output bit rate remains the same. In non-uniform sampling, we have to encode the sampling time points so that the average bit rate is the sum of the average amplitude quantization bit rate and the time quantization bit rate. After reconstructing the non-uniform samples onto a uniform time grid, we can evaluate the equivalent uniform sampling bit rate according to the SQNR of the reconstructed signal, which is the bit rate required when uniform sampling is applied in order to reach the same level of SQNR. Therefore, the performance metric of different random sampling architectures is the ratio between the equivalent uniform sampling bit rate and the total bit rate. According to this

metric, the SAR ADC random sampling architecture performs the best as it reduces the bit rate requirement by a factor of 6.5 compares with uniform sampling. That is to say, we can roughly reduce the power consumption in the ADC by a factor of 6.5 when using non-uniform sampling. However, the actual total power saving is not likely to be that significant as the the recovery algorithm used in the digital end consumes extra power.

Table 7: A performance comparison of the proposed random sampling architectures

	SAR ADC	ramp ADC	LC ADC
average sampling frequency (Hz)	12.29	38.35	119.57
amplitude quantization bit rate (b/s)	136.88	115.81	268.57
time quantization bit rate (b/s)	49.08	63.11	45.07
total bit rate (b/s)	185.95	178.92	313.64
SQNR (dB)	28.72	18.35	19.11
equivalent uniform sampling bit rate(b/s)	1192.59	762.02	793.68

Chapter IV

CONCLUSIONS AND FUTURE WORK DIRECTION

Sampling is the first and probably the most important step to preserve, retrieve and extract the information embedded in the signals that we are interested in. The Shannon-Nyquist sampling theorem provided a theoretical bound as how fast we should uniformly sample the continuous signal so that no information is lost in the sampling process. This thesis, however is dedicated to the exploration of the ways to break through the minimum sampling frequency by applying non-uniform sampling techniques.

The first undertaking was to study parallelism in conjunction with the uniform sampling process, which leads to an architecture known as time interleaving. The channel mismatches in the TIADC system make the system an instance of a recurrent non-uniform sampling system whose non-uniformities are detrimental to the performance of the system and need to be calibrated. Accordingly, we design a flexible and efficient architecture in Section 2.2 to compensate for the non-uniformities or mismatches in the TIADC system. The calibration architecture and a calibration algorithm have been established. Simulations have been conducted for the timing mismatches calibration effect of the proposed algorithm. As a key building block in the calibration architecture, the design of Farrow structured adjustable fractional delay filters has been investigated in detail in Section 2.3.1 to 2.3.4. A new modified Farrow structure is proposed in Section 2.3.5 to design adjustable FD filters for a given range of bandwidths and fractional delays. Detailed comparisons between the modified Farrow structure and other variants of Farrow structured FD filters are conducted afterwards. The Farrow structure is not limited to adjustable fractional delay

filters. It can also be used to implement adjustable lowpass, highpass and band-pass filters as well as adjustable multi-rate filters in Section 2.3.6 and 2.3.7 respectively. In Section 2.3.8, we further extend the Farrow structure to the design of adjustable polynomial phase responses.

Inspired by the theory of compressive sensing, another contribution of this thesis is to use randomization as a means to overcome the limit of the Nyquist rate. Section 3.1 investigate the impact of random sampling intervals or jitters on the power spectrum of the sampled signal. The theory shows that the aliases of the original signal can be well-shaped by choosing an appropriate probability distribution of the sampling intervals or jitters such that aliases can be reviews as a source of noise in the signal power spectrum. A new theoretical framework has been established in Section 3.2 to associate the probability mass function of the random sampling intervals or jitters with an aliasing shaping effect.

Based on the theoretical framework, the thesis proposes three random sampling architectures, i.e., SAR ADC, ramp ADC and level crossing ADC in Section 3.3.3 to 3.3.5, respectively, that can be easily implemented based on the corresponding standard ADC architectures. Detailed models and simulations are established to verify the effectiveness of the proposed architectures. A new reconstruction algorithm called the successive sine matching pursuit has also been proposed in Section 3.3.2 to recover a class of spectrally sparse signals from a sparse set of non-uniform samples onto a denser uniform time grid so that classic signal processing techniques can be applied afterwards.

The future work will extend the sparsity assumption in the spectral domain to a broader domain so that a wider class of signals can be sampled using the established random sampling theory.

REFERENCES

[1] Baraniuk, R., "Compressive sensing [lecture notes]," *Signal Processing Magazine, IEEE*, vol. 24, pp. 118 –121, July 2007.

[2] Berger, C., Wang, Z., Huang, J., and Zhou, S., "Application of compressive sensing to sparse channel estimation," *Communications Magazine, IEEE*, vol. 48, pp. 164 –174, Nov. 2010.

[3] Beutler, F. J., "Error-free recovery of signals from irregularly spaced samples," *SIAM Review*, vol. 8, no. 3, pp. pp. 328–335, 1966.

[4] Beutler, F. J. and Leneman, O. A. Z., "Random sampling of random process: Stationary point processes," *Information and Control*, vol. 9, pp. 325–344, 1966.

[5] Beutler, F. J. and Leneman, O. A. Z., "The theory of stationary point processes," *Acta Math.*, vol. 116, pp. 159–197, 1966.

[6] Beutler, F. J. and Leneman, O. A. Z., "The spectral analysis of impulse processes," *Information and Control*, vol. 12, pp. 236–258, 1968.

[7] Bilinskis, I., *Digital Alias-Free Signal Processing*. John Wiley & Sons, Ltd., 2007.

[8] Bilinskis, I. and Mikelsons, A., *Randomized Signal Processing*. Prentice-Hall International (UK), Ltd., 1992.

[9] Bresler, Y., "Spectrum-blind sampling and compressive sensing for continuous-index signals," in *Information Theory and Applications Workshop, 2008*, pp. 547 –554, Jan. 2008.

[10] Brown, J., J., "Multi-channel sampling of low-pass signals," *Circuits and Systems, IEEE Transactions on*, vol. 28, pp. 101 – 106, Feb. 1981.

[11] Brown, J.L., J. and Cabrera, S., "On well-posedness of the papoulis generalized sampling expansion," *Circuits and Systems, IEEE Transactions on*, vol. 38, pp. 554 –556, May 1991.

[12] Candes, E., Romberg, J., and Tao, T., "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *Information Theory, IEEE Transactions on*, vol. 52, pp. 489 – 509, Feb. 2006.

[13] Candes, E. and Wakin, M., "An introduction to compressive sampling," *Signal Processing Magazine, IEEE*, vol. 25, pp. 21 –30, Mar. 2008.

[14] Cenker, C., Feichtinger, H., and Herrmann, M., "Iterative algorithms in irregular sampling: a first comparison of methods," in *Computers and Communications, 1991. Conference Proceedings., Tenth Annual International Phoenix Conference on*, Mar. 1991.

[15] Chen, J. and Huo, X., "Theoretical results on sparse representations of multiple-measurement vectors," *Signal Processing, IEEE Transactions on*, vol. 54, pp. 4634–4643, Dec. 2006.

[16] Cheung, K. and Marks, R., "Iii-posed sampling theorems," *Circuits and Systems, IEEE Transactions on*, vol. 32, pp. 481 – 484, May 1985.

[17] Cotter, S., Rao, B., Engan, K., and Kreutz-Delgado, K., "Sparse solutions to linear inverse problems with multiple measurement vectors," *Signal Processing, IEEE Transactions on*, vol. 53, pp. 2477 – 2488, Jul. 2005.

[18] Dyer, K., Fu, D., Lewis, S., and Hurst, P., "An analog background calibration technique for time-interleaved analog-to-digital converters," *Solid-State Circuits, IEEE Journal of*, vol. 33, pp. 1912 –1919, Dec. 1998.

[19] Elbornsson, J. and Eklund, J.-E., "Blind estimation of timing errors in interleaved ad converters," in *Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on*, vol. 6, pp. 3913 –3916, 2001.

[20] Elbornsson, J., Gustafsson, F., and Eklund, J.-E., "Blind adaptive equalization of mismatch errors in a time-interleaved a/d converter system," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 51, pp. 151 – 158, Jan. 2004.

[21] Elbornsson, J., Gustafsson, F., and Eklund, J.-E., "Blind equalization of time errors in a time-interleaved adc system," *Signal Processing, IEEE Transactions on*, vol. 53, pp. 1413 – 1424, Apr. 2005.

[22] Eldar, Y. and Oppenheim, A., "Filterbank reconstruction of bandlimited signals from nonuniform and generalized samples," *Signal Processing, IEEE Transactions on*, vol. 48, pp. 2864 –2875, Oct. 2000.

[23] Farrow, C. W., "A continuously variable digital delay element," in *IEEE International Symposium on Circuits and Systems*, 1988.

[24] Fu, D., Dyer, K., Lewis, S., and Hurst, P., "A digital background calibration technique for time-interleaved analog-to-digital converters," *Solid-State Circuits, IEEE Journal of*, vol. 33, pp. 1904 –1911, Dec. 1998.

[25] Gardner, F., "Interpolation in digital modems. i. fundamentals," *Communications, IEEE Transactions on*, vol. 41, pp. 501 –507, Mar. 1993.

[26] Gilbert, A. C., Guha, S., Indyk, P., Muthukrishnan, S., and Strauss, M., "Near-optimal sparse fourier representations via sampling," in *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, STOC '02, (New York, NY, USA), pp. 152–161, ACM, 2002.

[27] Gurbuz, A., McClellan, J., and Cevher, V., "A compressive beamforming method," in *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pp. 2617 –2620, Apr. 2008.

[28] Gurbuz, A., McClellan, J., Scott, W., and Larson, G., "Seismic imaging and detection of underground tunnels," in *Signal Processing and Communications Applications, 2006 IEEE 14th*, pp. 1 –4, Apr. 2006.

[29] Hajji, R., Beanregard, F., and Ghannouchi, F., "Multitone power and intermodulation load-pull characterization of microwave transistors suitable for linear sspa's design," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 45, pp. 1093 –1099, Jul. 1997.

[30] Higgins, J., "A sampling theorem for irregularly spaced sample points (corresp.)," *Information Theory, IEEE Transactions on*, vol. 22, pp. 621 – 622, Sep. 1976.

[31] Jerri, A., "The shannon sampling theorem, its various extensions and applications: A tutorial review," *Proceedings of the IEEE*, vol. 65, pp. 1565 – 1596, Nov. 1977.

[32] Jin, H. and Lee, E., "A digital technique for reducing clock jitter effects in time-interleaved a/d converter," in *Circuits and Systems, 1999. ISCAS '99. Proceedings of the 1999 IEEE International Symposium on*, vol. 2, pp. 330 –333, July 1999.

[33] Jin, H. and Lee, E., "A digital-background calibration technique for minimizing timing-error effects in time-interleaved adcs," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 47, pp. 603 –613, July 2000.

[34] Johansson, H. and Gustafsson, O., "Linear-phase fir interpolation, decimation, and mth-band filters utilizing the farrow structure," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 52, pp. 2197 – 2207, Oct. 2005.

[35] Johansson, H. and Lowenborg, P., "Reconstruction of nonuniformly sampled bandlimited signals by means of digital fractional delay filters," *Signal Processing, IEEE Transactions on*, vol. 50, pp. 2757 – 2767, Nov. 2002.

[36] Kadec, M. I., "The exact value of the paley-wiener constant," *Soviet Math. Dokl.*, vol. 5, p. 559, 1964.

[37] Karam, L. and McClellan, J., "Complex chebyshev approximation for fir filter design," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 42, pp. 207 –216, Mar. 1995.

[38] Karam, L. and McClellan, J., "Design of optimal digital fir filters with arbitrary magnitude and phase responses," in *Circuits and Systems, 1996 IEEE International Symposium on*, vol. 2, pp. 385 –388, May 1996.

[39] Kester, W., "Data conversion handbook." Website, 2005. `http://www.analog.com/library/analogdialogue/archives/39-06/data\_conversion\_handbook.html`.

[40] Khoini-Poorfard, R. and Johns, D., "Time-interleaved oversampling convertors," *Electronics Letters*, vol. 29, pp. 1673 –1674, Sept. 1993.

[41] Khoini-Poorfard, R., Lim, L., and Johns, D., "Time-interleaved oversampling a/d converters: theory and practice," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 44, pp. 634 –645, Aug 1997.

[42] Kim, S.-J., Koh, K., Lustig, M., Boyd, S., and Gorinevsky, D., "An interior-point method for large-scale l1-regularized least squares," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, pp. 606 –617, Dec. 2007.

[43] Kroon, P. and Atal, B., "On the use of pitch predictors with high temporal resolution," *Signal Processing, IEEE Transactions on*, vol. 39, pp. 733 –735, Mar. 1991.

[44] Laakso, T., Valimaki, V., Karjalainen, M., and Laine, U., "Splitting the unit delay," *Signal Processing Magazine, IEEE*, vol. 13, pp. 30 –60, Jan. 1996.

[45] Lelandais-Perrault, C., Petrescu, T., Poulton, D., Duhamel, P., and Oksman, J., "Wideband, bandpass, and versatile hybrid filter bank a/d conversion for software radio," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 56, pp. 1772 –1782, Aug. 2009.

[46] Leneman, O. A. Z., "Random sampling of random process: Impulse processes," *Information and Control*, vol. 9, pp. 347–363, 1966.

[47] Löwenborg, P. and Johansson, H., "Minimax design of adjustable-bandwidth linear-phase fir filters," *Circuits And Systems Part I, IEEE Transactions on*, vol. 53, no. 2, pp. 431–439, 2006.

[48] Lustig, M., Donoho, D., Santos, J., and Pauly, J., "Compressed sensing mri," *Signal Processing Magazine, IEEE*, vol. 25, pp. 72 –82, Mar. 2008.

[49] Marvasti, F., Analoui, M., and Gamshadzahi, M., "Recovery of signals from nonuniform samples using iterative methods," *Signal Processing, IEEE Transactions on*, vol. 39, pp. 872 –878, Apr. 1991.

[50] Marvasti, F. and Lee, T., "Analysis and recovery of sample-and-hold and linearly interpolated signals with irregular samples," *Signal Processing, IEEE Transactions on*, vol. 40, pp. 1884 –1891, Aug. 1992.

[51] Mishali, M. and Eldar, Y., "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, pp. 375 –391, Apr. 2010.

[52] Mishali, M., Eldar, Y., and Elron, A., "Xampling: Signal acquisition and processing in union of subspaces," *Signal Processing, IEEE Transactions on*, vol. 59, pp. 4719 –4734, Oct. 2011.

[53] Mishali, M., Elron, A., and Eldar, Y., "Sub-nyquist processing with the modulated wideband converter," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 3626 –3629, Mar. 2010.

[54] Needell, D., Tropp, J., and Vershynin, R., "Greedy signal recovery review," in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, pp. 1048 –1050, Oct. 2008.

[55] Needell, D. and Tropp, J. A., "CoSaMP: Iterative signal recovery from incomplete and inaccurate samples," *Applied and Computational Harmonic Analysis*, vol. 26, pp. 301–321, Apr. 2008.

[56] Nguyen, H. and Do, M., "Hybrid filter banks with fractional delays: Minimax design and application to multichannel sampling," *Signal Processing, IEEE Transactions on*, vol. 56, pp. 3180 –3190, July 2008.

[57] Olsson, M., Johansson, H., and Lowenborg, P., "Simultaneous estimation of gain, delay, and offset utilizing the farrow structure," in *Circuit Theory and Design, 2007. ECCTD 2007. 18th European Conference on*, pp. 244 –247, Aug. 2007.

[58] Papoulis, A., "Generalized sampling expansion," *Circuits and Systems, IEEE Transactions on*, vol. 24, pp. 652 – 654, Nov. 1977.

[59] Pedro, J. and Carvalho, N., "Designing band-pass multisine excitations for microwave behavioral model identification," in *Microwave Symposium Digest, 2004 IEEE MTT-S International*, vol. 2, pp. 791 – 794, Jun. 2004.

[60] Pedro, J. and De Carvalho, N., "On the use of multitone techniques for assessing rf components' intermodulation distortion," *Microwave Theory and Techniques, IEEE Transactions on*, vol. 47, pp. 2393 – 2402, Dec. 1999.

[61] Pereira, J., Girao, P., and Serra, A., "An fft-based method to evaluate and compensate gain and offset errors of interleaved adc systems," *Instrumentation and Measurement, IEEE Transactions on*, vol. 53, pp. 423 – 430, Apr. 2004.

[62] Petraglia, A. and Mitra, S., "Analysis of mismatch effects among a/d converters in a time-interleaved waveform digitizer," *Instrumentation and Measurement, IEEE Transactions on*, vol. 40, pp. 831 –835, Oct. 1991.

[63] Petraglia, A. and Mitra, S., "High-speed a/d conversion incorporating a qmf bank," *Instrumentation and Measurement, IEEE Transactions on*, vol. 41, pp. 427 –431, Jun. 1992.

[64] Ping, F. and Bresler, Y., "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Acoustics, Speech, and Signal Processing, 1996. ICASSP-96. Conference Proceedings., 1996 IEEE International Conference on*, vol. 3, pp. 1688–1691, May 1996.

[65] Rudelson, M. and Vershynin, R., "Sparse reconstruction by convex relaxation: Fourier and gaussian measurements," *Information Sciences and Systems, 2006 40th Annual Conference on*, pp. 207 –212, Mar. 2006.

[66] Samadi, S., Ahmad, M., and Swamy, M., "Results on maximally flat fractional-delay systems," *Circuits and Systems I: Regular Papers, IEEE Transactions on*, vol. 51, pp. 2271 – 2286, Nov. 2004.

[67] Schulte, R. F., Henning, A., Tsao, J., Boesiger, P., and Pruessmann, K. P., "Design of broadband rf pulses with polynomial-phase response.," *Journal of Magnetic Resonance*, vol. 186, no. 2, pp. 167–175, 2007.

[68] Schulte, R. F., Tsao, J., Boesiger, P., and Pruessmann, K. P., "Equiripple design of quadratic-phase rf pulses.," *Journal of Magnetic Resonance*, vol. 166, no. 1, pp. 111–122, 2004.

[69] Shannon, C., "Communication in the presence of noise," *Proceedings of the IEEE*, vol. 72, pp. 1192 – 1201, Sept. 1984.

[70] Shapiro, H. S. and Silverman, R. A., "Aliasing-free sampling of random noises," *SIAM J. Appl. Math.*, vol. 8, pp. 225–236, 1960.

[71] Shu, H., Chen, T., and Francis, B., "Minimax design of hybrid multirate filter banks," *Circuits and Systems II: Analog and Digital Signal Processing, IEEE Transactions on*, vol. 44, pp. 120 –128, Feb. 1997.

[72] Simon, G. and Schoukens, J., "Robust broadband periodic excitation design," *Instrumentation and Measurement, IEEE Transactions on*, vol. 49, pp. 270 –274, Apr. 2000.

[73] Stoica, P., Li, J., and He, H., "Spectral analysis of nonuniformly sampled data: A new approach versus the periodogram," *Signal Processing, IEEE Transactions on*, vol. 57, pp. 843 –858, Mar. 2009.

[74] Tropp, J., "Random filters for compressive sampling," in *Information Sciences and Systems, 2006 40th Annual Conference on*, pp. 216 –217, Mar. 2006.

[75] Tropp, J. and Gilbert, A., "Signal recovery from random measurements via orthogonal matching pursuit," *Information Theory, IEEE Transactions on*, vol. 53, pp. 4655 –4666, Dec. 2007.

[76] Tropp, J., Laska, J., Duarte, M., Romberg, J., and Baraniuk, R., "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *Information Theory, IEEE Transactions on*, vol. 56, pp. 520 –544, Jan. 2010.

[77] Turner, L. R., "Inverse of the vandemonde matrix with applications," tech. rep.

[78] Unser, M., "Sampling-50 years after shannon," *Proceedings of the IEEE*, vol. 88, pp. 569 –587, Apr. 2000.

[79] Unser, M., Aldroubi, A., and Eden, M., "B-spline signal processing. i. theory," *Signal Processing, IEEE Transactions on*, vol. 41, pp. 821 –833, Feb. 1993.

[80] Unser, M., Aldroubi, A., and Eden, M., "B-spline signal processing. ii. efficiency design and applications," *Signal Processing, IEEE Transactions on*, vol. 41, pp. 834 –848, Feb. 1993.

[81] Vaidyanathan, P., "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial," *Proceedings of the IEEE*, vol. 78, pp. 56 –93, Jan. 1990.

[82] Vaidyanathan, P., *Multirate Systems And Filter Banks*. Prentice Hall Signal Processing Series, Pearson Education, 1993.

[83] Valimaki, V. and Karjalainen, M., "Implementation of fractional delay waveguide models using allpass filters," in *Acoustics, Speech, and Signal Processing. ICASSP-95., 1995 International Conference on*, vol. 2, pp. 1524 –1527, May 1995.

[84] Van Moer, W. and Rolain, Y., "A multisine based calibration for broadband measurements," in *Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE*, pp. 1 –6, May 2007.

[85] Van Moer, W., Rolain, Y., and Geens, A., "Measurement based nonlinear modeling of spectral regrowth," in *Microwave Symposium Digest. 2000 IEEE MTT-S International*, vol. 3, pp. 1467 –1470, 2000.

[86] Velazquez, S., Nguyen, T., Broadstone, S., and Roberge, J., "A hybrid filter bank approach to analog-to-digital conversion," in *Time-Frequency and Time-Scale Analysis, 1994., Proceedings of the IEEE-SP International Symposium on*, pp. 116 –119, Oct. 1994.

[87] Venkataramani, R. and Bresler, Y., "Perfect reconstruction formulas and bounds on aliasing error in sub-nyquist nonuniform sampling of multiband signals," *Information Theory, IEEE Transactions on*, vol. 46, pp. 2173 –2183, Sep. 2000.

[88] Vogel, C., Draxelmayr, D., and Kuttner, F., "Compensation of timing mismatches in time-interleaved analog-to-digital converters through transfer characteristics tuning," in *Circuits and Systems, 2004. MWSCAS '04. The 2004 47th Midwest Symposium on*, vol. 1, pp. I – 341–4 vol.1, July 2004.

[89] Vogel, C. and Kubin, G., "Analysis and compensation of nonlinearity mismatches in time-interleaved adc arrays," in *Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on*, vol. 1, pp. 593–596, May 2004.

[90] Vogel, C. and Kubin, G., "Time-interleaved adcs in the context of hybrid filter banks," in *Proceedings of the 2004 URSI International Symposium on Signals, Systems, and Electronics (ISSSE 2004)*, pp. 217–217, Aug. 2004.

[91] Vogel, C. and Kubin, G., "Modeling of time-interleaved adcs with nonlinear hybrid filter banks," *AEU International Journal of Electronics and Communications*, vol. 59, no. 5, pp. 288–296, 2005.

[92] Vogel, C., *Modeling, Identification, and Compensation of Channel Mismatch Errors in Time-Interleaved Analog-to-Digital Converters*. PhD thesis, Graz University of Technology, Austria, 2005.

[93] Wingham, D., "The reconstruction of a band-limited function and its fourier transform from a finite number of samples at arbitrary locations by singular value decomposition," *Signal Processing, IEEE Transactions on*, vol. 40, pp. 559 –570, Mar. 1992.

[94] Yen, J., "On nonuniform sampling of bandwidth-limited signals," *Circuit Theory, IRE Transactions on*, vol. 3, pp. 251 – 257, Dec. 1956.

VITA

Chenchi Luo was born in Shanghai, China in November 1983. He received his B.S. degree in Electrical Engineering and M.S. degree in Automation of Electrical Power Systems, both from Shanghai Jiao Tong University in 2006 and 2008, respectively. He also received a M.S. degree in Electrical and Computer Engineering from Georgia Institute of Technology in 2009 and is expected to receive his Ph.D. degree in fall 2012. He is currently a Ph.D. student at the Center for Signal and Information Processing, Georgia institute of technology under the supervision of Dr. James H. McClellan. He has a rich industrial working experience with DSP algorithm design and software implementation. He was employed by National Instruments, Shanghai, China, from 2007 to 2008 developing the LabVIEW virtual instrumentation software. During his graduate program, he was employed by Texas Instruments as a research intern in the System and Applications R&D Center, Dallas, TX in 2010 and 2011. His research interests include algorithm and software development for DSP based systems, compressive sensing, digital filter design and optimization, non-uniform sampling and reconstruction, digital image and audio processing, statistical and adaptive signal processing.